

OPI DATE 14/03/96 APPLN. ID 34972/95
AOJP DATE 02/05/96 PCT NUMBER PCT/US95/10918



AU9534972

7

| | | | |
|---|--|-----------|--|
| (51) International Patent Classification ⁶ : A61K 39/395, C12N 5/10, 1/21, 15/63, C12Q 1/68, C07K 14/00, 16/18, C07H 21/04 | | A1 | (11) International Publication Number: WO 96/05861 (43) International Publication Date: 29 February 1996 (29.02.96) |
| (21) International Application Number: PCT/US95/10918 (22) International Filing Date: 23 August 1995 (23.08.95) (30) Priority Data: 294,522 23 August 1994 (23.08.94) US 470,868 6 June 1995 (06.06.95) US (71) Applicant: MILLENIUM PHARMACEUTICALS, INC. [US/US]; 5th floor, 640 Memorial Drive, Cambridge, MA 02139 (US). (72) Inventor: TARTAGLIA, Louis, A.; 87 Spruce Street, Water- stown, MA 02172 (US). (74) Agents: CORUZZI, Laura, A. et al.; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US). | | | (81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i> |

(54) Title: COMPOSITIONS AND METHODS FOR THE TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

(57) Abstract

The present invention relates to methods and compositions for the treatment of body weight disorders, including, but not limited to, obesity. Specifically, the present invention identifies and describes genes which are differentially expressed in body weight disorder states, relative to their expression in normal, or non-body weight disorder states, and/or in response to manipulations relevant to appetite and/or weight regulation. Further, the present invention identifies and describes genes via the ability of their gene products to interact with gene products involved in body weight disorders and/or appetite and/or body weight regulation. Still further, the present invention provides methods for the identification and therapeutic use of compounds as treatments of body weight disorders. Additionally, the present invention describes methods for the diagnostic evaluation and prognosis of various body weight disorders, and for the identification of subjects exhibiting a predisposition to such conditions.

COMPOSITIONS AND METHODS FOR THE TREATMENT OF
BODY WEIGHT DISORDERS, INCLUDING OBESITY

This is a continuation-in-part of Serial No.

5 08/470,868, filed June 6, 1995, which is a
continuation-in-part of Serial No. 08/294,522, filed
August 23, 1994, both applications of which are
incorporated herein by reference in their entirety.

1. INTRODUCTION

10 The present invention relates to methods and
compositions for the treatment of body weight
disorders, including, but not limited to, obesity.
Specifically, the present invention identifies and
15 describes genes which are differentially expressed in
body weight disorder states, relative to their
expression in normal, or non-body weight disorder
states, and also identifies genes which are
differentially expressed in response to manipulations
20 relevant to appetite and/or weight regulation.
Further, the present invention identifies and
describes genes via the ability of their gene products
to interact with gene products involved in body weight
disorders and/or to interact with gene products which
25 are relevant to appetite and/or body weight
regulation. Still further, the present invention
provides methods for the identification and
therapeutic use of compounds as treatments of body
weight disorders. Additionally, the present invention
30 describes methods for the diagnostic evaluation and
prognosis of various body weight disorders, and for
the identification of subjects exhibiting a
predisposition to such conditions.

2. BACKGROUND OF THE INVENTION

35 Body weight disorders, including eating
disorders, represent major health problems in all

industrialized countries. Obesity, the most prevalent of eating disorders, for example, is the most important nutritional disorder in the western world, with estimates of its prevalence ranging from 30% to 50% within the middle-aged population. Other body weight disorders, such as anorexia nervosa and bulimia nervosa which together affect approximately 0.2% of the female population of the western world, also pose serious health threats. Further, such disorders as anorexia and cachexia (wasting) are also prominent features of other diseases such as cancer, cystic fibrosis, and AIDS.

Obesity, defined as an excess of body fat relative to lean body mass, also contributes to other diseases. For example, this disorder is responsible for increased incidences of diseases such as coronary artery disease, stroke, and diabetes. Obesity is not merely a behavioral problem, i.e., the result of voluntary hyperphagia. Rather, the differential body composition observed between obese and normal subjects results from differences in both metabolism and neurologic/metabolic interactions. These differences seem to be, to some extent, due to differences in gene expression, and/or level of gene products or activity. The nature, however, of the genetic factors which control body composition are unknown, and attempts to identify molecules involved in such control have generally been empiric and the parameters of body composition and/or substrate flux are monitored have not yet been identified (Friedman, J.M. et al., 1991, Mammalian Gene 1:130-144).

The epidemiology of obesity strongly shows that the disorder exhibits inherited characteristics, (Stunkard, 1990, N. Eng. J. Med. 322:1483). Moll et al., have reported that, in many populations, obesity

seems to be controlled by a few genetic loci, (Moll et al. 1991, Am. J. Hum. Gen. 49:1243). In addition, human twin studies strongly suggest a substantial genetic basis in the control of body weight, with
5 estimates of heritability of 80-90% (Simopoulos, A.P. & Childs B., eds., 1989, in "Genetic Variation and Nutrition in Obesity", World Review of Nutrition and Diabetes 63, S. Karger, Basel, Switzerland; Borjeson, M., 1976, Acta. Paediatr. Scand. 65:279-287).

10 Further, studies of non-obese persons who deliberately attempted to gain weight by systematically over-eating were found to be more resistant to such weight gain and able to maintain an elevated weight only by very high caloric intake. In
15 contrast, spontaneously obese individuals are able to maintain their status with normal or only moderately elevated caloric intake.

In addition, it is a commonplace experience in animal husbandry that different strains of swine,
20 cattle, etc., have different predispositions to obesity. Studies of the genetics of human obesity and of models of animal obesity demonstrate that obesity results from complex defective regulation of both food intake, food induced energy expenditure and of the
25 balance between lipid and lean body anabolism.

There are a number of genetic diseases in man and other species which feature obesity among their more prominent symptoms, along with, frequently, dysmorphic features and mental retardation. Although no
30 mammalian gene associated with an obesity syndrome has yet been characterized in molecular terms, a number of such diseases exist in humans. For example, Prader-Willi syndrome (PWS) affects approximately 1 in 20,000 live births, and involves poor neonatal muscle tone,
35 facial and genital deformities, and generally obesity.

The genetics of PWS are very complex, involving, for example, genetic imprinting, in which development of the disease seems to depend upon which parent contributes the abnormal PWS allele. In approximately
5 half of all PWS patients, however, a deletion on the long arm of chromosome 11 is visible, making the imprinting aspect of the disease difficult to reconcile. Given the various symptoms generated, it seems likely that the PWS gene product may be required
10 for normal brain function, and may, therefore, not be directly involved in adipose tissue metabolism.

In addition to PWS, many other pleiotropic syndromes which include obesity as a symptom have been characterized. These syndromes are more genetically
15 straightforward, and appear to involve autosomal recessive alleles. The diseases, which include, among others, Ahlstroem, Carpenter, Bardet-Biedl, Cohen, and Morgagni-Stewart-Monel Syndromes.

Animals having mutations which lead to syndromes
20 that include obesity symptoms have also been identified. Attempts have been made to utilize such animals as models for the study of obesity. The best studied animal models for genetic obesity are mice which contain the autosomal recessive mutations ob/ob
25 (obese) and db/db (diabetes). These mutations are on chromosomes 6 and 4, respectively, but lead to clinically similar pictures of obesity, evident starting at about 1 month of age, which include hyperphagia, severe abnormalities in glucose and
30 insulin-metabolism, very poor thermo-regulation and non-shivering thermogenesis, and extreme torpor and underdevelopment of the lean body mass. Restriction of the diet of these animals to restore a more normal body fat mass to lean body mass ration is fatal and
35 does not result in a normal habitus.

Although the phenotypes of db/db and ob/ob mice are similar, the lesions are distinguishable by means of parabiosis. The feeding of normal mice and, putatively, all mammals, is regulated by satiety factors. The ob/ob mice are apparently unable to express the satiety factor, while the db/db mouse is unresponsive to it.

In addition to ob and db, several other single gene mutations resulting in obesity in mice have been identified. These include the yellow mutation at the agouti locus, which causes a pleiotropic syndrome which causes moderate adult onset obesity, a yellow coat color, and a high incidence of tumor formation (Herberg, L. and Coleman, D.L., 1977, Metabolism 26:59), and an abnormal anatomic distribution of body fat (Coleman, D.L., 1978, Diabetologia 14:141-148). Additionally, mutations at the fat and tubby loci cause moderately severe, maturity-onset obesity with somewhat milder abnormalities in glucose homeostasis than are observed in ob and db mice (Coleman, D.L., and Eicher, E.M., 1990, J. Heredity 81:424-427). Further, autosomal dominant mutations at the adipose locus of chromosome 7, have been shown to cause obesity.

Other animal models include fa/fa (fatty) rats, which bear many similarities to the ob/ob and db/db mice, discussed above. One difference is that, while fa/fa rats are very sensitive to cold, their capacity for non-shivering thermogenesis is normal. Torpor seems to play a larger part in the maintenance of obesity in fa/fa rats than in the mice mutants. In addition, inbred mouse strains such as NZO mice and Japanese KK mice are moderately obese. Certain hybrid mice, such as the Wellesley mouse, become spontaneously fat. Further, several desert rodents,

such as the spiny mouse, do not become obese in their natural habitats, but do become so when fed on standard laboratory feed.

Animals which have been used as models for obesity have also been developed via physical or pharmacological methods. For example, bilateral lesions in the ventromedial hypothalamus (VMH) and ventrolateral hypothalamus (VLH) in the rat are associated, respectively, with hyperphagia and gross obesity and with aphagia and cachexia. Further, it has been demonstrated that feeding monosodium-glutamate (MSG) to new born mice also results in an obesity syndrome.

Attempts have been made to utilize such animal models in the study molecular causes of obesity. For example, adipsin, a murine serine protease with activity closely similar to human complement factor D, produced by adipocytes, has been found to be suppressed in ob/ob, db/db and MSG-induced obesity (Flier, 1987, Science 237:405). The suppression of adipsin precedes the onset of obesity in each model (Lowell, 1990, Endocrinology 126:1514). Further studies have mapped the locus of the defect in these models to activity of the adipsin promoter (Platt, 1989, Proc. Natl. Acad. Sci. USA 86:7490). Further, alterations have been found in the expression of neuro-transmitter peptides in the hypothalamus of the ob/ob mouse (Wilding, 1993, Endocrinology 132:1939), of glucose transporter proteins in islet β -cells (Ohneda, 1993, Diabetes 42:1065) and of the levels of G-proteins (McFarlane-Anderson, 1992, Biochem. J. 282:15).

To date however, no gene, in humans, has been found which is causative in the processes leading to obesity. Given the severity and prevalence, however,

of disorders, including obesity, which affect body weight and body composition, there exists a great need for the systematic identification of such body weight disorder-causing genes.

5

3. SUMMARY OF THE INVENTION

The present invention relates to methods and compositions for the treatment of body weight disorders, including, but not limited to, obesity. Specifically, the present invention identifies and describes genes which are differentially expressed in body weight disorder states, relative to their expression in normal, or non-body weight disorder states, and also identifies genes which are differentially expressed in response to manipulations relevant to appetite and/or body weight regulation. Such differentially expressed genes may represent "target genes" and/or "fingerprint genes". Further, the present invention identifies and describes genes, termed "pathway genes", via the ability of their gene products to interact with gene products involved in body weight disorders and/or to interact with gene products which are relevant to appetite and body weight regulation. Pathway genes may also exhibit target gene and/or fingerprint gene characteristics.

"Differential expression", as used herein, refers to both quantitative as well as qualitative differences in the genes' temporal and/or tissue expression patterns. "Fingerprint gene," as used herein, refers to a differentially expressed gene whose expression pattern may be utilized as part of a prognostic or diagnostic body weight disorder evaluation, or which, alternatively, may be used in methods for identifying compounds useful for the treatment of body weight disorders. "Target gene", as

used herein, refers to a differentially expressed gene involved in body weight disorders and/or appetite or body regulation such that modulation of the level of target gene expression or of target gene product activity may act to ameliorate symptoms of body weight disorders including, but are not limited to, obesity.

This invention is based, in part on systematic, search strategies involving body weight disorder experimental paradigms coupled with sensitive gene expression assays.

The present invention also describes the products of such fingerprint, target, and pathway genes, describes antibodies to such gene products, and still further describes cell- and animal-based models of body weight disorders to which such gene products may contribute.

The invention further provides methods for the identification of compounds which modulate the expression of genes or the activity of gene products involved in body weight disorders and processes relevant to appetite and/or body weight regulation. Still further, the present invention describes methods for the treatment of body weight disorders which may involve the administration of such compounds to individuals exhibiting body weight disorder symptoms or tendencies.

Additionally, the present invention describes methods for prognostic and diagnostic evaluation of various body weight disorders, and for the identification of subjects exhibiting a predisposition to such disorders.

The Examples presented in Sections 6-10, below, demonstrate the successful use of the body weight disorder paradigms of the invention to identify body weight disorder target genes.

4. DESCRIPTION OF THE FIGURES

FIG. 1. Differential display comparing RNAs from liver tissue of lean and obese mice. Each group (1-11) of four lanes shows the pattern obtained with one primer pair combination, for a total of eleven different primer pair combinations. All lanes are products of a polymerase chain reaction (PCR) in which T₁₁GG was used as the 3' oligonucleotide and one of eleven different arbitrary 10mer oligonucleotides was used as the 5' oligonucleotide. Within each group of four lanes, the loading is as follows, from left to right: C57B1/6J lean control (marked "C"); C57B1/6J ob/ob (marked "ob"); C57B1/Ks lean control (marked "C"); and C57B1/Ks db/db (marked "db"). An arrow indicates a band (designated L36) that is differential between obese and lean samples amplified by the same primer pair, specifically, primer pair 6.

FIG. 2. Northern blot analysis confirming differential regulation of a gene corresponding to band L36. Poly A⁺ RNA (1μg/lane) obtained from the original liver total RNA preparations was hybridized with a cDNA probe prepared by random priming of reamplified lane L36 (see materials and methods, below, in Section 6.1). Lane 1, C57B1/6J lean control ("C"); lane 2, C57B1/6J ob/ob ("ob"); lane 3 C57B1/Ks lean control ("C"); and lane 4, C57B1/Ks db/db ("db").

FIG. 3A. Consensus nucleotide sequence of L36 amplified band (SEQ. ID NO: 1). The IUPAC-IUB Standard Code is used (with the addition of the "X" designation, as shown below) in this nucleotide sequence and those nucleotide sequences listed in the figures which follow. Upper case letters refer to perfect consensus matches at a particular base pair

position. Lower case letters refer to base pair positions at which there was a less than perfect consensus match. Specifically, the code used was as follows:

| | | | |
|----|-------------|-------------------------------|----------------|
| 5 | | | |
| | <u>Code</u> | <u>Base</u> | <u>Meaning</u> |
| | A | A | Adenine |
| | C | C | Cytosine |
| | G | G | Guanine |
| | T | T | Thymine |
| 10 | U | U | Uracil (RNA) |
| | R | A or G | Purine |
| | Y | C or T (or U) | Pyrimidine |
| | K | G or T (or U) | Keto |
| | M | A or C | Amino |
| | S | G or C | Strong |
| 15 | W | A or T (or U) | Weak |
| | B | C, G, T (or U) | not A |
| | D | G, A, T (or U) | not C |
| | H | A, C, T (or U) | not G |
| | V | A, C, or G | not T |
| 20 | N | A, C, G, T (or U) | Any |
| | X | A, C, G, T (or U), or none | Any or none |

FIG. 3B. Alignment of L36 consensus nucleotide sequence with a mouse stearyl-CoA desaturase nucleotide sequence (SEQ. ID NOS:2, 3). It will be noted that there are two alignments listed for this L36/mouse stearyl-CoA desaturase match. Each represents a highly statistically significant alignment, and together, these alignments represent very highly significant matches. This is the case for each of the matches listed in the figures, below, which have greater than one alignment listed.

FIG. 4. Alignment of P3 consensus nucleotide sequence (SEQ ID NO.:4, 39) with a mouse glutamine synthetase nucleotide sequence (SEQ ID NOS.:5, 6).

5 FIG. 5. Alignment of P13 consensus nucleotide sequence (SEQ ID NO.:7, 40, 41, 42) with a mouse islet regenerating protein nucleotide sequence (SEQ NOS.:8, 9, 10, 11).

10 FIG. 6. Alignment of F5 consensus nucleotide sequence (SEQ ID NO.:12) with a mouse alpha-amylase nucleotide sequence (SEQ ID NO.:13).

15 FIG. 7. Alignment of murine C5 consensus nucleotide sequence (SEQ ID NO.:14) with a rabbit uncoupling protein nucleotide sequence (SEQ ID NO.:15).

20 FIG. 8. Alignment of L31/F74 consensus nucleotide sequence (SEQ ID NO.:16) with a mouse major urinary protein II nucleotide sequence (SEQ ID NO.:17).

25 FIG. 9. Alignment of L7/L21 consensus nucleotide sequence (SEQ ID NO.:18) with a mouse cytochrome oxidase c subunit I nucleotide sequence (SEQ ID NO.:19).

30 FIG. 10. Alignment of L29 consensus nucleotide sequence (SEQ ID NO.:20) with a mouse testosterone 15-alpha hydroxylase nucleotide sequence (SEQ ID NO.:21).

35 FIG. 11. Alignment of L38 consensus nucleotide sequence (SEQ ID NO.:22, 43) with a mouse 24p3 (a

lipocalin family member of unknown function)
nucleotide sequence (SEQ ID NO.:23, 24).

5 FIG. 12. Alignment of L37 consensus nucleotide
sequence (SEQ ID NO.:25, 44, 45) with a mouse p6-5 (a
mouse sequence 86% homologous to rat preproelastase I)
nucleotide sequence (SEQ ID NO.:26, 27, 28).

10 FIG. 13. Alignment of L57 consensus nucleotide
sequence (SEQ ID NO.:29, 46, 47, 48) with a mouse
orphan hormone receptor nucleotide sequence (SEQ ID
NO.:30, 31, 32, 33).

15 FIG. 14. Full length F49 cDNA clone. A cDNA
putatively encoding the full length coding sequence of
the F49 gene was isolated and its nucleotide sequence
is listed herein (SEQ ID NO.:34). The F49 coding
sequence encodes a 96 amino acid protein whose
sequence is also listed herein (SEQ ID NO.:35). The
20 initiating methionine codon and the termination codon
are boxed.

25 FIG. 15. Hydropathy plot of the F49 gene
product.

30 FIG. 16. Depicted herein are, first, the full
length mouse C5 nucleotide sequence (SEQ ID NO.:36),
and second, the amino acid sequence (SEQ ID NO.:37)
encoded by the mouse C5 gene. The initiating
methionine and the termination codon are boxed.

35 FIG. 17. Human C5 cDNA. The bottom line of the
Figure depicts the nucleotide sequence of a human C5
cDNA clone (SEQ ID NO.:38); the top line of the Figure
depicts the portion of the human C5 amino acid

sequence (SEQ ID NO: 56) derived from the cDNA nucleotide sequence.

FIG. 18. Full length C5 amino acid sequence (SEQ ID NO: 51).

FIG. 19. Northern analysis of tissue distribution of human C5 mRNA.

FIG. 20. Alignment of H27 consensus nucleotide sequence (SEQ ID NO: 49) with a mouse autoantigen La nucleotide sequence (SEQ ID NO: 50).

FIG. 21. Alignment of F84 consensus nucleotide sequence (SEQ ID NO: 52, 54) with a mouse cytochrome p450 IID nucleotide sequence (SEQ ID NO: 53, 55).

FIG. 22. Mouse L34 cDNA nucleotide sequence (SEQ ID NO: 57).

5. DETAILED DESCRIPTION OF THE INVENTION

Methods and compositions for the treatment of body weight disorders, including, but not limited to, obesity. Genes, termed "target genes" and/or "fingerprint genes", are described which are differentially expressed in body weight disorder states, relative to their expression in normal, or non-body weight disorder states, and/or which are differentially expressed in response to manipulations relevant to appetite and/or body weight regulation. Additionally, genes, termed "pathway genes", are described whose gene products exhibit an ability to interact with gene products involved in body weight disorders and/or with gene products which are relevant to appetite and/or body weight regulation. Methods

for the identification of such fingerprint, target, and pathway genes are also described.

Further, the gene products of such fingerprint, target, and pathway genes are described, antibodies to such gene products are described, as are cell- and animal-based models of body weight disorders to which such gene products may contribute.

Described, below, are methods for the identification of compounds which modulate the expression of genes or the activity of gene products involved in body weight disorders and processes relevant to appetite and/or body weight regulation. Additionally described, below, are methods for the treatment of body weight disorders.

Also discussed, below, are methods for prognostic and diagnostic evaluation of various body weight disorders, and for the identification of subjects exhibiting a predisposition to such disorders.

20

5.1. IDENTIFICATION OF DIFFERENTIALLY EXPRESSED AND PATHWAY GENES

Described herein are methods for the identification of genes which are involved in body weight disorder states, and/or which are involved in appetite and body weight regulation. Such genes may represent genes which are differentially expressed in body weight disorder states relative to their expression in normal, or non-body weight disorder states. Further, such genes may represent genes which are differentially regulated in response to manipulations relevant to appetite and body weight regulation. Such differentially expressed genes may represent "target" and/or "fingerprint" genes. Methods for the identification of such differentially

expressed genes are described, below, in Section 5.1.1. Methods for the further characterization of such differentially expressed genes, and for their identification as target and/or fingerprint genes, are presented, below, in Section 5.1.3.

In addition, methods are described herein, in Section 5.1.2, for the identification of genes, termed "pathway genes", involved in body weight disorder states, and/or in appetite or body weight regulation. "Pathway gene", as used herein, refers to a gene whose gene product exhibits the ability to interact with gene products involved in body weight disorders and/or to interact with gene products which are relevant to appetite or body weight regulation. A pathway gene may be differentially expressed and, therefore, may have the characteristics of a target and/or fingerprint gene.

"Differential expression" as used herein refers to both quantitative as well as qualitative differences in the genes' temporal and/or tissue expression patterns. Thus, a differentially expressed gene may qualitatively have its expression activated or completely inactivated in normal versus body weight disorder states, or under control versus experimental conditions. Such a qualitatively regulated gene will exhibit an expression pattern within a given tissue or cell type which is detectable in either control or body weight disorder subjects, but is not detectable in both. Alternatively, such a qualitatively regulated gene will exhibit an ~~expression pattern~~ within a given tissue or cell type which is detectable in either control or experimental subjects, but is not detectable in both. "Detectable", as used herein, refers to an RNA expression pattern which is detectable via the standard techniques of differential

display, RT-PCR and/or Northern analyses, which are well known to those of skill in the art.

Alternatively, a differentially expressed gene may have its expression modulated, *i.e.*, quantitatively increased or decreased, in normal versus body weight disorder states, or under control versus experimental conditions. The degree to which expression differs in normal versus body weight disorder or control versus experimental states need only be large enough to be visualized via standard characterization techniques, such as, for example, the differential display technique described below. Other such standard characterization techniques include but are not limited to, quantitative RT (reverse transcriptase) PCR and Northern analyses.

Differentially expressed genes may be further described as target genes and/or fingerprint genes. "Fingerprint gene," as used herein, refers to a differentially expressed gene whose expression pattern may be utilized as part of a prognostic or diagnostic body weight disorder evaluation, or which, alternatively, may be used in methods for identifying compounds useful for the treatment of body weight disorders. A fingerprint gene may also have the characteristics of a target gene or a pathway gene.

"Target gene", as used herein, refers to a differentially expressed gene involved in body weight disorders and/or appetite or body weight regulation in a manner by which modulation of the level of target gene expression or of target gene product activity may act to ameliorate symptoms of body weight disorders including, but are not limited to, obesity. A target gene may also have the characteristics of a fingerprint gene and/or a pathway gene.

5.1.1. METHODS FOR THE IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENES

5 A variety of methods may be utilized for the identification of genes which are involved in body weight disorder states, and/or which are involved in appetite and body weight regulation. Described in Section 5.1.1.1 are several experimental paradigms which may be utilized for the generation of subjects and samples which may be used for the identification of such genes. Material from the paradigm control and experimental subjects may be characterized for the presence of differentially expressed gene sequences as discussed, below, in Section 5.1.1.2.

15 5.1.1.1. PARADIGMS FOR THE IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENES

Among the paradigms which may be utilized for the identification of differentially expressed genes involved in, for example, body weight disorders, are paradigms designed to analyze those genes which may be involved in short term appetite control. Accordingly, such paradigms are referred to as "short term appetite control paradigms." These paradigms may serve to identify genes involved in signalling hunger and satiety.

25 In one embodiment of such a paradigm, test subjects, preferably mice, may be fed normally prior to the initiation of the paradigm study, then divided into one control and two experimental groups. The control group would then be maintained on ad lib nourishment, while the first experimental group ("fasted group") would be fasted, and the second experimental group ("fasted-refed group") would initially be fasted, and would then be offered a highly palatable meal shortly before the collection of tissue samples. Each test animal should be weighted

35

immediately prior to and immediately after the experiment. The Example presented in Section 7, below, demonstrates the use of such short term appetite paradigms to identify gene sequences which are differentially expressed in control versus fasting and versus refed animals.

Among additional paradigms which may be utilized for the identification of differentially expressed genes involved in, for example, body weight disorders, are paradigms designed to analyze those genes which may be involved genetic obesity. Accordingly, such paradigms are referred to as "genetic obesity paradigms". In the case of mice, for example, such paradigms may identify genes regulated by the ob, db, and/or tub gene products. In the case of rats, for example, such paradigms may identify genes regulated by the fatty (fa) gene product.

In one embodiment of such a paradigm, test subjects may include ob/ob, db/db, and/or tub/tub experimental mice and lean littermate control animals. Such animals would be offered normal nourishment for a given period, after which tissue samples would be collected for analysis. The Examples presented in Sections 6 and 8, below, demonstrate the use of such genetic obesity paradigms in identifying gene sequences which are differentially expressed in obese versus lean animals.

In additional embodiments, ob/ob, db/db, and/or tub/tub experimental mice and lean control animals may be utilized as part of the short term appetite control paradigms discussed above, or as part of the set point and/or drug study paradigms discussed below.

Paradigms which may be utilized for the identification of differentially expressed genes involved in body weight disorders may include

paradigms designed to identify those genes which may be regulated in response to changes in body weight. Such paradigms may be referred to as "set point paradigms".

5 In one embodiment of such a paradigm, test subjects, preferably mice, may be fed normally prior to the initiation of the paradigm study, then divided into one control and two experimental groups. The control group would then be maintained on an ad lib
10 diet of normal nourishment in order to calculate daily food intake. The first experimental group ("underweight group") would then be underfed by receiving some fraction of normal food intake, 60-90% of normal, for example, so as to reduce and maintain
15 the group's body weight to some percentage, for example 80%, of the control group. The second experimental group ("overweight group") would be overfed by receiving a diet which would bring the group to some level above that of the control, for
20 example 125% of the control group. Tissue samples would then be obtained for analysis. The Example presented in Section 9, below, demonstrates the use of such set point paradigms to identify gene sequences which are differentially expressed in control versus
25 overweight and/or underweight conditions.

 Additionally, human subjects may be utilized for the identification of obesity-associated genes. In one embodiment of such a paradigm, tissue samples may be obtained from obese and lean human subjects and
30 analyzed for the presence of genes which are differentially expressed in the tissue of one group as opposed to another (e.g. differentially expressed in lean versus obese subjects). In another embodiment, obese human subjects may be studied over the course of
35 a period of weight loss, achieved through food

restriction. Tissue from these previously obese subjects may be analyzed for differential expression of gene products relative to tissue obtained from control (lean, non-previously obese) and obese subjects.

Paradigms may be utilized for the identification of differentially expressed genes involved in body weight disorders may additionally include paradigms designed to identify genes associated with body weight disorders induced by some physical manipulation to the test subject, such as, for example, hypothalamic lesion-induced body weight disorders. For example, bilateral lesions in the ventromedial hypothalamus (VMH) of rodents may be utilized to induce hyperphagia and gross obesity in test subjects, while bilateral lesions in the ventrolateral hypothalamus (VLH) of rodents may be utilized to induce aphagia in test subjects. In such paradigms, tissue from hypothalamic-lesioned test subjects and from control subjects would be analyzed for the identification of genes which are differentially expressed in control versus lesioned animals.

Drugs known to affect (e.g., ameliorate) human or animal body weight and/or appetite (such as short term appetite) may be incorporated into paradigms designed to identify genes which are involved in body weight disorders and/or body weight or appetite regulation. Accordingly, such paradigms are referred to as "drug study paradigms". Such compounds may include known therapeutics, as well as compounds that are not useful as therapeutics due to, for example, their harmful side effects. Among the categories of control and test subjects which may be utilized in such paradigms are, for example, lean subjects, obese subjects, and obese subjects which have received the drug of

interest. In various embodiments of the paradigms, subjects such as these may be fed a normal ad lib diet, a caloric restriction maintained diet, or a caloric restriction ad lib diet. Control and test
5 subjects may additionally be paired i.e., the control and test subjects may be fed via a coupled feeding device such that both control and test subjects receive identical amounts and types of food).

10 5.1.1.2. ANALYSIS OF PARADIGM MATERIAL

In order to identify differentially expressed genes, RNA, either total or mRNA, may be isolated from one or more tissues of the subjects utilized in paradigms such as those described, above, in Section
15 5.1.1. RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique which does not select against the isolation of mRNA may be utilized for the purification of such RNA
20 samples. See, for example, Ausubel, F.M. et al., eds., 1987-1993, Current Protocols in Molecular Biology, John Wiley & Sons, Inc. New York, which is incorporated herein by reference in its entirety. Additionally, large numbers of tissue samples may
25 readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, P. (1989, U.S. Patent No. 4,843,155), which is incorporated herein by reference in its entirety.

30 Transcripts within the collected ~~RNA samples~~ which represent RNA produced by differentially expressed genes may be identified by utilizing a variety of methods which are well known to those of skill in the art. For example, differential screening
35 (Tedder, T.F. et al., 1988, Proc. Natl. Acad. Sci. USA

85:208-212), subtractive hybridization (Hedrick, S.M. et al., 1984, Nature 308:149-153; Lee, S.W. et al., 1984, Proc. Natl. Acad. Sci. USA 88:2825), and, preferably, differential display (Liang, P. and Pardee, A.B., 1992, Science 257:967-971; U.S. Patent No. 5,262,311, which is incorporated herein by reference in its entirety), may be utilized to identify nucleic acid sequences derived from genes that are differentially expressed.

10 Differential screening involves the duplicate screening of a cDNA library in which one copy of the library is screened with a total cell cDNA probe corresponding to the mRNA population of one cell type while a duplicate copy of the cDNA library is screened
15 with a total cDNA probe corresponding to the mRNA population of a second cell type. For example, one cDNA probe may correspond to a total cell cDNA probe of a cell type or tissue derived from a control subject, while the second cDNA probe may correspond to
20 a total cell cDNA probe of the same cell type or tissue derived from an experimental subject. Those clones which hybridize to one probe but not to the other potentially represent clones derived from genes differentially expressed in the cell type of interest
25 in control versus experimental subjects.

Subtractive hybridization techniques generally involve the isolation of mRNA taken from two different sources, e.g., control and experimental tissue or cell type, the hybridization of the mRNA or single-stranded
30 cDNA reverse-transcribed from the isolated mRNA, and the removal of all hybridized, and therefore double-stranded, sequences. The remaining non-hybridized, single-stranded cDNAs, potentially represent clones derived from genes that are differentially expressed
35 in the two mRNA sources. Such single-stranded cDNAs

are then used as the starting material for the construction of a library comprising clones derived from differentially expressed genes.

The differential display technique describes a procedure, utilizing the well known polymerase chain reaction (PCR; the experimental embodiment set forth in Mullis, K.B., 1987, U.S. Patent No. 4,683,202) which allows for the identification of sequences derived from genes which are differentially expressed. First, isolated RNA is reverse-transcribed into single-stranded cDNA, utilizing standard techniques which are well known to those of skill in the art. Primers for the reverse transcriptase reaction may include, but are not limited to, oligo dT-containing primers, preferably of the 3' primer type of oligonucleotide described below.

Next, this technique uses pairs of PCR primers, as described below, which allow for the amplification of clones representing a random subset of the RNA transcripts present within any given cell. Utilizing different pairs of primers allows each of the mRNA transcripts present in a cell to be amplified. Among such amplified transcripts may be identified those which have been produced from differentially expressed genes.

The 3' oligonucleotide primer of the primer pairs may contain an oligo dT stretch of 10-13, preferably 11, dT nucleotides at its 5' end, which hybridizes to the poly(A) tail of mRNA or to the complement of a cDNA reverse transcribed from an mRNA poly(A) tail. Second, in order to increase the specificity of the 3' primer, the primer may contain one or more, preferably two, additional nucleotides at its 3' end. Because, statistically, only a subset of the mRNA derived sequences present in the sample of interest will

hybridize to such primers, the additional nucleotides allow the primers to amplify only a subset of the mRNA derived sequences present in the sample of interest. This is preferred in that it allows more accurate and complete visualization and characterization of each of the bands representing amplified sequences.

The 5' primer may contain a nucleotide sequence expected, statistically, to have the ability to hybridize to cDNA sequences derived from the tissues of interest. The nucleotide sequence may be an arbitrary one, and the length of the 5' oligonucleotide primer may range from about 9 to about 15 nucleotides, with about 13 nucleotides being preferred.

Arbitrary primer sequences cause the lengths of the amplified partial cDNAs produced to be variable, thus allowing different clones to be separated by using standard denaturing sequencing gel electrophoresis.

PCR reaction conditions should be chosen which optimize amplified product yield and specificity, and, additionally, produce amplified products of lengths which may be resolved utilizing standard gel electrophoresis techniques. Such reaction conditions are well known to those of skill in the art, and important reaction parameters include, for example, length and nucleotide sequence of oligonucleotide primers as discussed above, and annealing and elongation step temperatures and reaction times.

The pattern of clones resulting from the reverse transcription and amplification of the mRNA of two different cell types is displayed via sequencing gel electrophoresis and compared. Differentially expressed genes are indicated by differences in the two banding patterns.

Once potentially differentially expressed gene sequences have been identified via bulk techniques such as, for example, those described above, the differential expression of such putatively

- 5 differentially expressed genes should be corroborated. Corroboration may be accomplished via, for example, such well known techniques as Northern analysis, quantitative RT PCR or RNase protection.

- 10 Upon corroboration, the differentially expressed genes may be further characterized, and may be identified as target and/or fingerprint genes, as discussed, below, in Section 5.1.3.

- Also, amplified sequences of differentially expressed genes obtained through, for example, differential display may be used to isolate full length clones of the corresponding gene. The full length coding portion of the gene may readily be isolated, without undue experimentation, by molecular biological techniques well known in the art. For example, the isolated differentially expressed amplified fragment may be labeled and used to screen a cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

- 25 PCR technology may also be utilized to isolate full length cDNA sequences. As described, above, in this Section, the isolated, amplified gene fragments obtained through differential display have 5' terminal ends at some random point within the gene and usually have 3' terminal ends at a position corresponding to the 3' end of the transcribed portion of the gene. Once nucleotide sequence information from an amplified fragment is obtained, the remainder of the gene (i.e., the 5' end of the gene, when utilizing differential display) may be obtained using, for example, RT-PCR.

35

In one embodiment of such a procedure for the identification and cloning of full length gene sequences, RNA may be isolated, following standard procedures, from an appropriate tissue or cellular source. A reverse transcription reaction may then be performed on the RNA using an oligonucleotide primer complimentary to the mRNA that corresponds to the amplified fragment, for the priming of first strand synthesis. Because the primer is anti-parallel to the mRNA, extension will proceed toward the 5' end of the mRNA. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Using the two primers, the 5' portion of the gene is amplified using PCR. Sequences obtained may then be isolated and recombined with previously isolated sequences to generate a full-length cDNA of the differentially expressed genes of the invention. For a review of cloning strategies and recombinant DNA techniques, see e.g., Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; and Ausubel et al., 1989, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

5.1.2. METHODS FOR THE IDENTIFICATION OF PATHWAY GENES

Methods are described herein for the identification of pathway genes. "Pathway gene", as used herein, refers to a gene whose gene product exhibits the ability to interact with gene products involved in body weight disorders and/or to interact with gene products which are relevant to appetite or body weight regulation. A pathway gene may be differentially expressed and, therefore, may have the characteristics of a target and/or fingerprint gene.

Any method suitable for detecting protein-protein interactions may be employed for identifying pathway gene products by identifying interactions between gene products and gene products known to be involved in body weight disorders and/or involved in appetite or body regulation. Such known gene products may be cellular or extracellular proteins. Those gene products which interact with such known gene products represent pathway gene products and the genes which encode them represent pathway genes.

Among the traditional methods which may be employed are co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns. Utilizing procedures such as these allows for the identification of pathway gene products. Once identified, a pathway gene product may be used, in conjunction with standard techniques, to identify its corresponding pathway gene. For example, at least a portion of the amino acid sequence of the pathway gene product may be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique (see, e.g., Creighton, 1983, "Proteins: Structures and Molecular Principles", W.H. Freeman & Co., N.Y., pp.34-49). The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for pathway gene sequences. Screening made be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well-known. (See, e.g., Ausubel, supra., and PCR Protocols: A Guide to Methods and Applications, 1990, Innis, M. et al., eds. Academic Press, Inc., New York).

35

Additionally, methods may be employed which result in the simultaneous identification of pathway genes which encode the protein interacting with a protein involved in body weight disorder states and/or appetite and body weight regulation. These methods include, for example, probing expression libraries with labeled protein known or suggested to be involved in body weight disorders and/or appetite or body weight regulation, using this protein in a manner similar to the well known technique of antibody probing of λ gt11 libraries.

One method which detects protein interactions in vivo, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version of this system has been described (Chien et al., 1991, Proc. Natl. Acad. Sci. USA, 88:9578-9582) and is commercially available from Clontech (Palo Alto, CA).

Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one consists of the DNA-binding domain of a transcription activator protein fused to a known protein, in this case, a protein known to be involved in body weight disorders and or processes relevant to appetite and/or weight regulation, and the other consists of the transcription activator protein's activation domain fused to an unknown protein that is encoded by a cDNA which has been recombined into this plasmid as part of a cDNA library. The plasmids are transformed into a strain of the yeast Saccharomyces cerevisiae that contains a reporter gene (e.g., lacZ) whose regulatory region contains the transcription activator's binding sites. Either hybrid protein alone cannot activate transcription of the reporter gene: the DNA-binding domain hybrid cannot because it does not provide

activation function and the activation domain hybrid cannot because it cannot localize to the activator's binding sites. Interaction of the two hybrid proteins reconstitutes the functional activator protein and
5 results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodology may be used to screen activation domain libraries for proteins that interact with a known "bait" gene
10 product. By way of example, and not by way of limitation, gene products known to be involved in body weight disorders and/or appetite or body weight regulation may be used as the bait gene products. These include but are not limited to the intracellular
15 domain of receptors for such hormones as neuropeptide Y, galanin, interostatin, insulin, and CCK. Total genomic or cDNA sequences are fused to the DNA encoding an activation domain. This library and a plasmid encoding a hybrid of the bait gene product
20 fused to the DNA-binding domain are cotransformed into a yeast reporter strain, and the resulting transformants are screened for those that express the reporter gene. For example, and not by way of limitation, the bait gene can be cloned into a vector
25 such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein. These colonies are purified and the library plasmids responsible for reporter gene expression are isolated. DNA sequencing is then used to identify the proteins
30 encoded by the library plasmids.

A cDNA library of the cell line from which proteins that interact with bait gene product are to be detected can be made using methods routinely practiced in the art. According to the particular
35 system described herein, for example, the cDNA

fragments can be inserted into a vector such that they are translationally fused to the activation domain of GAL4. This library can be co-transformed along with the bait gene-GAL4 fusion plasmid into a yeast strain which contains a lacZ gene driven by a promoter which contains GAL4 activation sequence. A cDNA encoded protein, fused to GAL4 activation domain, that interacts with bait gene product will reconstitute an active GAL4 protein and thereby drive expression of the lacZ gene. Colonies which express lacZ can be detected by their blue color in the presence of X-gal. The cDNA can then be purified from these strains, and used to produce and isolate the bait gene-interacting protein using techniques routinely practiced in the art.

Once a pathway gene has been identified and isolated, it may be further characterized as, for example, discussed below, in Section 5.1.3.

5.1.3. CHARACTERIZATION OF DIFFERENTIALLY EXPRESSED AND PATHWAY GENES

Differentially expressed genes, such as those identified via the methods discussed, above, in Section 5.1.1, and pathway genes, such as those identified via the methods discussed, above, in Section 5.1.2, above, as well as genes identified by alternative means, may be further characterized by utilizing, for example, methods such as those discussed herein. Such genes will be referred to herein as "identified genes".

Analyses such as those described herein, yield information regarding the biological function of the identified genes. An assessment of the biological function of the differentially expressed genes, in addition, will allow for their designation as target and/or fingerprint genes.

- Specifically, any of the differentially expressed genes whose further characterization indicates that a modulation of the gene's expression or a modulation of the gene product's activity may ameliorate any of the body weight disorders of interest will be designated "target genes", as defined, above, in Section 5.1. Such target genes and target gene products, along with those discussed below, will constitute the focus of the compound discovery strategies discussed, below, in Section 5.3. Further, such target genes, target gene products and/or modulating compounds can be used as part of the body weight disorder treatment methods described, below, in Section 5.4.
- Any of the differentially expressed genes whose further characterization indicates that such modulations may not positively affect body weight disorders of interest, but whose expression pattern contributes to a gene expression "fingerprint" pattern correlative of, for example, a body weight disorder state will be designated a "fingerprint gene".
- "Fingerprint patterns" will be more fully discussed below, in Section 5.7.1. It should be noted that each of the target genes may also function as fingerprint genes, as well as may all or a portion of the pathway genes.
- It should further be noted that the pathway genes may also be characterized according to techniques such as those described herein. Those pathway genes which yield information indicating that they are differentially expressed and that modulation of the gene's expression or a modulation of the gene product's activity may ameliorate any of the body weight disorders of interest will be also be designated "target genes". Such target genes and target gene products, along with those discussed

above, will constitute the focus of the compound discovery strategies discussed, below, in Section 5.3 and can be used as part of the treatment methods described in Section 5.4, below.

5 It should be additionally noted that the characterization of one or more of the pathway genes may reveal a lack of differential expression, but evidence that modulation of the gene's activity or expression may, nonetheless, ameliorate body weight
10 disorder symptoms. In such cases, these genes and gene products would also be considered a focus of the compound discovery strategies of Section 5.3, below.

 In instances wherein a pathway gene's characterization indicates that modulation of gene
15 expression or gene product activity may not positively affect body weight disorders of interest, but whose expression is differentially expressed and contributes to a gene expression fingerprint pattern correlative of, for example, a body weight disorder
20 state, such pathway genes may additionally be designated as fingerprint genes.

 A variety of techniques can be utilized to further characterize the identified genes. First, the nucleotide sequence of the identified genes, which may
25 be obtained by utilizing standard techniques well known to those of skill in the art, may, for example, be used to reveal homologies to one or more known sequence motifs which may yield information regarding the biological function of the identified gene
30 product.

 Second, an analysis of the tissue and/or cell type distribution of the mRNA produced by the identified genes may be conducted, utilizing standard techniques well known to those of skill in the art.
35 Such techniques may include, for example, Northern,

- 5 RNase protection and RT-PCR analyses. Such analyses provide information as to, for example, whether the identified genes are expressed in tissues or cell types expected to contribute to the body weight disorders of interest. Such analyses may also provide quantitative information regarding steady state mRNA regulation, yielding data concerning which of the identified genes exhibits a high level of regulation in, preferably, tissues which may be expected to contribute to the body weight disorders of interest.
- 10 Additionally, standard in situ hybridization techniques may be utilized to provide information regarding which cells within a given tissue express the identified gene. Such an analysis may provide information regarding the biological function of an identified gene relative to a subset of the disorder in instances wherein only a subset of the cells within the tissue is thought to be relevant to the body weight disorder.
- 20 Third, the sequences of the identified genes may be used, utilizing standard techniques, to place the genes onto genetic maps, e.g., mouse (Copeland, N.G. and Jenkins, N.A., 1991, Trends in Genetics 7:113-118) and human genetic maps (Cohen, D., et al., 1993, Nature 366:698-701). Such mapping information may yield information regarding the genes' importance to human disease by, for example, identifying genes which map within genetic regions to which known genetic body weight disorders map.
- 30 Fourth, the biological function of the identified genes may be more directly assessed by utilizing relevant *in vivo* and *in vitro* systems. *In vivo* systems may include, but are not limited to, animal systems which naturally exhibit body weight disorder-like symptoms, or ones which have been engineered to
- 35

exhibit such symptoms. Further, such systems may include systems for the further characterization of body weight disorders, and/or appetite or body weight regulation, and may include, but are not limited to, naturally occurring and transgenic animal systems such as those described, above, in Section 5.1.1.1, and Section 5.2.4.1, below. In vitro systems may include, but are not limited to, cell-based systems comprising cell types known or suspected of contributing to the body weight disorder of interest. Such cells may be wild type cells, or may be non-wild type cells containing modifications known to, or suspected of, contributing to the body weight disorder of interest. Such systems are discussed in detail, below, in Section 5.2.4.2.

In further characterizing the biological function of the identified genes, the expression of these genes may be modulated within the in vivo and/or in vitro systems, i.e., either overexpressed or underexpressed in, for example, transgenic animals and/or cell lines, and its subsequent effect on the system then assayed. Alternatively, the activity of the product of the identified gene may be modulated by either increasing or decreasing the level of activity in the in vivo and/or in vitro system of interest, and its subsequent effect then assayed.

The information obtained through such characterizations may suggest relevant methods for the treatment of body weight disorders involving the gene of interest. Further, relevant methods for the control of appetite and body weight regulation involving the gene of interest may be suggested by information obtained from such characterizations. For example, treatment may include a modulation of gene expression and/or gene product activity.

Characterization procedures such as those described herein may indicate where such modulation should involve an increase or a decrease in the expression or activity of the gene or gene product of interest.

- 5 Such methods of treatment are discussed, below, in Section 5.4.

5.2. DIFFERENTIALLY EXPRESSED AND PATHWAY GENES

- 10 Identified genes, which include, but are not limited to, differentially expressed genes such as those identified in Section 5.1.1, above, and pathway genes, such as those identified in Section 5.1.2, above, are described herein. Specifically, the nucleic acid sequences and gene products of such
15 identified genes are described. Further, antibodies directed against the identified genes' products, and cell- and animal-based models by which the identified genes may be further characterized and utilized are also discussed in this Section.

20

5.2.1. DIFFERENTIALLY EXPRESSED GENE SEQUENCES

- Differentially expressed nucleotide sequences are shown in FIGS. 3A, 4-14, 16-17 and 20-21. Table 1
25 lists differentially expressed genes (P3, P13, F5, F49, murine C5, human C5, L31/F74, L7/L21, L29, L38, L37, L57, H27, F84 and L34) identified through, for example, the paradigms discussed, above, in Section 5.1.1.1, and, below, in the examples presented in
30 Sections 6-10. Table 1 also summarizes information regarding the further characterization of such genes. Table 2 lists E. coli clones, deposited with the Agricultural Research Service Culture Collection (NRRL), which contain sequences found within the F49
35 and human C5 genes listed in Table 1.

In Table 1, the differential expression patterns revealed via, for example, one or more of the paradigm conditions described in Section 5.1.1.1, above, are summarized under the column headed "Paradigm Expression Pattern". For each of the tested genes, the paradigm which was used and the difference in the expression of the gene in experimental versus control tissues is shown. "↑" indicates that gene expression is increased (i.e., there is an increase in the amount of detectable mRNA produced by a given gene) in experimental versus control tissue or cell type, while "↓" indicates that gene expression is decreased (i.e., there is an decrease in the amount of detectable mRNA produced by a given gene) in experimental versus control tissue or cell type. Further, "+" indicates that gene expression is activated in experimental versus control tissue or cell type, i.e., mRNA is detectable in experimental tissue or cell type whereas none is detectable in control tissue or cell type, while "-" would indicate that gene expression is inactivated in experimental versus control tissue or cell type, i.e., while mRNA is detectable in control tissue or cell type, it is no longer detectable in experimental tissue or cell type. "Detectable" as used herein, refers to levels of mRNA which are detectable via standard differential display, Northern and/or RT-PCR techniques which are well known to those of skill in the art. "Increased" and "decreased", as used herein, refer to an increase or decrease, respectively in level of mRNA present in experimental versus control tissue or cell type which is detectable via standard differential display, Northern, and/or RT-PCT techniques which are well known to those of skill in the art.

35

Tissue expression patterns are also summarized in Table 1. The column headed "First Detection" indicates the first tissue or cell type in which differential expression of the gene was detected. The column headed "Tissue/Cell Dist." lists tissues and/or cell types in which expression of the gene has been tested and whether expression of the gene within a given tissue or cell type has been observed. Specifically, "+" indicates detectable mRNA from the gene of interest, while "-" refers to no detectable mRNA from the gene of interest. Unless otherwise noted, "+" and "-" refer to both control and experimental samples. "Detectable", as used herein, is as defined earlier in this Section.

Additionally, the physical locus to which the gene maps on the human and/or mouse chromosome map is indicated in the column headed "Locus". Further, in instances wherein the genes correspond to genes known to be found in nucleic acid databases, references (*i.e.*, citations and/or gene names) to such known genes are listed in the column headed "Ref".

The genes listed in Table 1 can be obtained using cloning methods well known to those of skill in the art, and which include, but are not limited to, the use of appropriate probes to detect the genes within an appropriate cDNA or gDNA (genomic DNA) library. (See, for example, Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Hanson Laboratories, which is incorporated herein by reference in its entirety.) Probes for the sequences reported herein can be obtained directly from the isolated clones deposited with the NRRL, as indicated in Table 2, below. Alternatively, oligonucleotide probes for the genes can be synthesized based on the DNA sequences disclosed herein in FIGs. 3A, 4-14, 16-

- 17 and 20-22. With respect to the previously reported genes, oligonucleotides can be synthesized or produced based on the sequences provided for the previously known genes described in the following references:
- 5 glutamine synthetase (P3): Bhandari et al., 1991, J. Biol. Chem. 266:7784-7792; islet regenerating protein (P13): Unno, M. et al., 1987, J. Biol. Chem. 268:15974-15982; Terazono, K. et al., 1988, J. Biol. Chem. 263:2111-2114; Watanabe, T., 1990, J. Biol.
- 10 Chem. 265:7432-7439; alpha amylase (F5): Schibler, U. et al., 1986, in "Oxford Surveys on Eukaryotic Genes", Maclean, N., ed. 3:210, Oxford Univ. Press, New York; Schibler et al., 1982, J. Mol. Biol. 155:247-266; mouse major urinary protein II (L31/F74): Shahan, K.
- 15 et al., 1987, Mol. Cell. Biol. 1:1938-1946; mouse cytochrome C oxidase Subunit I (L7/L21): Bibb et al., 1981, Cell 26:167-180; mouse testosterone 15-alpha hydroxylase (L15): Squires, E.J. and Negism:, M., 1988, J. Biol. Chem. 263:4166-4171; mouse 24p3 (L38):
- 20 Flower, D.R. et al., 1991, Biochem. Biophys. Res. Comm. 180:69-74; Hraba-Renevey, S. et al., 1989, Oncogene 4:601-608; mouse p6-5 (L37): Yamasaki, N. et al., 1987, Eur. J. Immunol. 17:247-253; mouse orphan Nuclear hormone Reception (L57): Forman et al., 1994,
- 25 Mol. Endocrinol. 8:1253-1261; autoantigen La (H27): Genbank Accession No. L00993; and mouse cytochrome p450 IID (F84): Matsunaga, E. et al., 1990, J. Mol. Evol. 30:155-169; mouse L34 (homolog of the human lrp 130 gene): Hou, J. et al., 1994, In Vitro Cell. Dev.
- 30 Biol. Anim. 30A:111-114.

The probes can be used to screen cDNA libraries prepared from an appropriate cell, cell line or tissue in which the gene is transcribed. Appropriate cell lines can include, for example, preadipocyte cell

35 lines such as 3T3-A1 and TA1 mouse preadipocyte cell

lines, liver cell lines, such as the Hepa1-6 mouse liver cell line and the HepG2 human liver cell line.

5

10

15

20

25

30

35

TABLE 1

Differentially Expressed Genes

| | | | | | | |
|----|--|-----------------------------------|--------------------|---|------------|-----|
| 5 | Gene | Paradigm Expression Pattern | First Detection | Tissue/ Cell Dist. | Locus | Ref |
| 10 | P3 (SEQ. ID NO:4, 39) | #Pancreas (fasted) | Pancreas | | | 1 |
| | P13 (SEQ. ID NO:7, 40-42) | #Pancreas (fasted) | Pancreas | | | 2 |
| | F5 (SEQ. ID NO:12) | #Adipose (fasted) | Adipose | | | 3 |
| 15 | F49 (SEQ. ID NO:34) | +Adipose (db/db) | Adipose | Adipose (-) db/db; (-) lean control) Muscle (-) Small Intestin e (-) Hypothal amus (-) Liver (-) Pancreas (-) | Chrom 2 | |
| 20 | | | | | | |
| 25 | | | | | | |
| 30 | murine C5 ¹⁰ (SEQ. ID NO: 36) | #Adipose (ob/ob and db/db) | Adipose | | | |

35

| | Gene | Paradigm Expression Pattern | First Detection | Tissue/ Cell Dist. | Locus | Ref |
|----|-----------------------------------|--|--------------------|--|-------|-----|
| 5 | L31/F74 (SEQ. ID NO:16) | ↓Liver and Adipose (ob/ob and db/db) ↓Liver (underweight) | Adipose; Liver | Liver (+) Adipose (+) Muscle (-) | | 4 |
| 10 | L7/L21 (SEQ. ID NO:18) | ↓Liver (fasted, ob/ob, and db/db) | Liver | | | 5 |
| 15 | L29 (SEQ. ID NO:20) | ↓Liver (ob/ob) | Liver | | | 6 |
| | L38 (SEQ. ID NO:22, 43) | ↓Liver (ob/ob and db/db) | Liver | | | 7 |
| | L37 (SEQ. NO:25, 44- 45) | ↓Liver (ob/ob) | Liver | | | 8 |
| 20 | L57 (SEQ. ID NO:29, 46- 48) | ↓Liver (underweight) | Liver | | | 9 |
| 25 | Human C5 (SEQ ID NO.:38) | | | Heart (+) Brain(-) Placenta (+) Lung (+) Liver (+) Muscle (+) Kidney (+) Pancreas (+) | | |
| 30 | | | | | | |
| 35 | | | | | | |

| | Gene | Paradigm Expression Pattern | First Detection | Tissue/Cell Dist. | Locus | Ref |
|----|-------------------------|-----------------------------|-----------------|-------------------|-------|-----|
| 5 | H27 (SEQ ID NO: 49) | Hypothalamus (fasted) | Hypothalamus | | | 11 |
| | F84 (SEQ ID NO: 52, 54) | Adipose (underweight) | Adipose | | | 12 |
| 10 | L34 (SEQ ID NO: 57) | Liver (ob/ob) | Liver | | | 13 |

- ¹Mouse glutamine synthetase: Bhandari et al., 1991, J. Biol. Chem. 266:7784-7792.
- ²Mouse islet regenerating protein: Unno, M. et al., 1993, J. Biol. Chem. 268:15974-15982.
- ³Mouse α -amylase: Schibler, U. et al., 1986, in "Oxford Surveys on Eukaryotic Genes", Maclean, N., ed., 3:210, Oxford Univ. Press, New York; Schibler et al., 1982, J. Mol. Biol. 155:247-266.
- ⁴Mouse major urinary protein II: Shaban, K. et al., 1987, Mol. Cell. Biol. 7:1938-1946.
- ⁵Mouse cytochrome C oxidase Subunit I: Raikhinstein, M. and Hankoglu, I., 1993, Proc. Natl. Acad. Sci. USA 90:10509-10513; Bibb et al., 1981, Cell 25:167-180.
- ⁶Mouse testosterone 15- α hydroxylase: Squires, E.J. and Negishi, M., 1988, J. Biol. Chem. 263:4166-4171.
- ⁷Mouse 24p3: Flower, D.R. et al., 1991, Biochem. Biophys. Res. Comm. 180:69-74; Hraba-Renevey, S. et al., 1989, Oncogene 4: 601-608.
- ⁸Mouse p6-5: Yamasaki, N. et al., 1987, Eur. J. Immunol. 17:247-253.
- ⁹Mouse orphan nuclear hormone receptor: Forman et al., 1994, Mol. Endocrinol. 8:1253-1261.
- ¹⁰The mouse C5 sequence was first identified via sequence homology. C5 was then subsequently tested in ob and db mice, at which time it was identified to represent a differentially expressed gene sequence.
- ¹¹Mouse autoantigen La: Genbank Accession No. L00993.
- ¹²Mouse cytochrome p450 IID: Matsunaga, E. et al., 1990, J. Mol. Evol. 30:155-169.
- ¹³The mouse L34 gene represents the mouse homolog of the human lrp 130 gene: Hou, J. et al., 1994, In Vitro Cell. Dev. Biol. Anim. 30A:111-114.

30 Table 2, below, lists isolated cDNA clones that contain genes listed in Table 1.

TABLE 2

| | GENE | cDNA CLONE |
|----|------|------------|
| 35 | F49 | famf049a |

| | |
|----------|----------|
| human C5 | fahs005a |
|----------|----------|

As used herein, "differentially expressed gene" (i.e. target and fingerprint gene) or "pathway gene" refers to (a) a gene containing: at least one of the DNA sequences disclosed herein (as shown in FIGS. 3A, 4-14, 16-17 and 20-22), or contained in the clones listed in Table 2, as deposited with the NRRL; (b) any DNA sequence that encodes the amino acid sequence encoded by: the DNA sequences disclosed herein (as shown in FIGS. 3A, 4-14, 16-17 and 20-22), contained in the clones listed in Table 2, as deposited with the NRRL, or contained within the coding region of the gene to which the DNA sequences disclosed herein (as shown in FIGS. 3A, 4-14, 16-17 and 20-22) or contained in the clones listed in Table 2, as deposited with the NRRL, belong; (c) any DNA sequence that hybridizes to the complement of: the coding sequences disclosed herein (as shown in FIGS. 3A, 4-14, 16-17 and 20-22), contained in clones listed in Table 2, as deposited with the NRRL, or contained within the coding region of the gene to which the DNA sequences disclosed herein (as shown in FIGS. 3A, 4-14, 16-17 and 20-22) or contained in the clones listed in Table 2, as deposited with the NRRL, belong, under highly stringent conditions, e.g., hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3), and encodes a gene product functionally equivalent to a gene product encoded by a gene of (a), above; and/or (d) any DNA sequence that hybridizes to the complement of: the

coding sequences disclosed herein, (as shown in FIGS. 3A, 4-14, 16-17 and 20-22) contained in the clones listed in Table 2, as deposited with the NRRL, or contained within the coding region of the gene to which DNA sequences disclosed herein (as shown in FIGS. 3A, 4-14, 16-17 and 20-22) or contained in the clones, listed in Table 2, as deposited with the NRRL, belong, under less stringent conditions, such as moderately stringent conditions, e.g., washing in 0.2xSSC/0.1% SDS at 42°C (Ausubel et al., 1989, supra), yet which still encodes a gene product, functionally equivalent to a gene product encoded by a gene of (a), above.

The invention also includes nucleic acid molecules, preferably DNA molecules, that hybridize to, and are therefore the complements of, the DNA sequences (a) through (d), in the preceding paragraph. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the nucleic acid molecules are deoxyoligonucleotides ("oligos"), highly stringent conditions may refer, e.g., to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may act as target gene antisense molecules, useful, for example, in target gene regulation and/or as antisense primers in amplification reactions of target, fingerprint, and/or pathway gene nucleic acid sequences. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for target gene regulation. Still further, such molecules may be used as components of diagnostic methods whereby the presence of, or predisposition to, a body weight disorder, may be detected.

The invention also encompasses (a) DNA vectors that contain any of the foregoing coding sequences and/or their complements (i.e., antisense); (b) DNA expression vectors that contain any of the foregoing coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences; and (c) genetically engineered host cells that contain any of the foregoing coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell. As used herein, regulatory elements include but are not limited to inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. The invention includes fragments of any of the DNA sequences disclosed herein.

In addition to the gene sequences described above, homologues of these gene sequences as may, for example, be present in other species, preferably human in instances wherein the above-described gene sequences are not human gene sequences, may be identified and isolated, without undue experimentation, by molecular biological techniques well known in the art. Further, there may exist genes at other genetic loci within the genome that encode proteins which have extensive homology to one or more domains of such gene products. These genes may also be identified via similar techniques.

For example, an ~~isolated differentially expressed~~ gene sequence may be labeled and used to screen a cDNA library constructed from mRNA obtained from the organism of interest. Hybridization conditions will be of a lower stringency when the cDNA library was derived from an organism different from the type of

organism from which the labeled sequence was derived. Alternatively, the labeled fragment may be used to screen a genomic library derived from the organism of interest, again, using appropriately stringent

5 conditions. Such low stringency conditions will be well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see,

10 for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Springs Harbor Press, N.Y.; and Ausubel et al., 1989, Current Protocols in Molecular Biology, (Green Publishing Associates and Wiley Interscience, N.Y.).

15 Further, a previously unknown differentially expressed or pathway gene-type sequence may be isolated by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences within the gene of interest. The

20 template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from human or non-human cell lines or tissue known or suspected to express a differentially expressed or pathway gene allele. The PCR product may be subcloned and

25 sequenced to insure that the amplified sequences represent the sequences of a differentially expressed or pathway gene-like nucleic acid sequence.

The PCR fragment may then be used to isolate a full length cDNA clone by a variety of methods. For

30 example, the amplified fragment may be used to screen a bacteriophage cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

PCR technology may also be utilized to isolate

35 full length cDNA sequences. For example, RNA may be

isolated, following standard procedures, from an appropriate cellular or tissue source. A reverse transcription reaction may be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Thus, cDNA sequences upstream of the amplified fragment may easily be isolated. For a review of cloning strategies which may be used, see e.g., Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Springs Harbor Press, N.Y.; and Ausubel et al., 1989, Current Protocols in Molecular Biology, (Green Publishing Associates and Wiley Interscience, N.Y.).

In cases where the differentially expressed or pathway gene identified is the normal, or wild type, gene, this gene may be used to isolate mutant alleles of the gene. Such an isolation is preferable in processes and disorders which are known or suspected to have a genetic basis. Mutant alleles may be isolated from individuals either known or suspected to have a genotype which contributes to body weight disorder symptoms. Mutant alleles and mutant allele products may then be utilized in the therapeutic and diagnostic assay systems described below.

A cDNA of the mutant gene may be isolated, for example, by using PCR, a technique which is well known to those of skill in the art. In this case, the first cDNA strand may be synthesized by hybridizing a oligo-dT oligonucleotide to mRNA isolated from tissue known to, or suspected of, being expressed in an individual putatively carrying the mutant allele, and by

extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5'- end of the normal gene. Using these two
5 primers, the product is then amplified via PCR, cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the
10 mutant gene to that of the normal gene, the mutation(s) responsible for the loss or alteration of function of the mutant gene product can be ascertained.

Alternatively, a genomic or cDNA library can be constructed and screened using DNA or RNA,
15 respectively, from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. The normal gene or any suitable fragment thereof may then be labeled and used as a probe to identify the
20 corresponding mutant allele in the library. The clone containing this gene may then be purified through methods routinely practiced in the art, and subjected to sequence analysis as described, above, in this Section..

25 Additionally, an expression library can be constructed utilizing DNA isolated from or cDNA synthesized from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. In
30 this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal gene product, as described, below, in Section 5.2.3. (For
35 screening techniques, see, for example, Harlow, E. and

Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.) In cases where the mutation results in an expressed gene product with altered function (e.g., as a result of a missense mutation), a polyclonal set of antibodies are likely to cross-react with the mutant gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis as described in this Section, above.

10

5.2.2. DIFFERENTIALLY EXPRESSED AND PATHWAY GENE PRODUCTS

Differentially expressed and pathway gene products include those proteins encoded by the differentially expressed and pathway gene sequences described in Section 5.2.1, above, as for example, the peptides listed in FIGS. 14 (SEQ ID NO: 34), 16 (SEQ ID NO: 36), 17 (SEQ. ID NO.: 56) and 18 (SEQ ID NO: 51).

20

In addition, differentially expressed and pathway gene products may include proteins that represent functionally equivalent gene products. Such an equivalent differentially expressed or pathway gene product may contain deletions, additions or substitutions of amino acid residues within the amino acid sequence encoded by the differentially expressed or pathway gene sequences described, above, in Section 5.2.1, but which result in a silent change, thus producing a functionally equivalent differentially expressed or pathway gene product. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, 'nonpolar (hydrophobic) amino acids include alanine, leucine,

25

30

35

isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Functionally equivalent", as utilized herein, refers to either a protein capable of exhibiting a substantially similar in vivo activity as the endogenous differentially expressed or pathway gene products encoded by the differentially expressed or pathway gene sequences described in Section 5.2.1, above. Alternatively, when utilized as part of assays such as those described, below, in Section 5.3, "functionally equivalent" may refer to peptides capable of interacting with other cellular or extracellular molecules in a manner substantially similar to the way in which the corresponding portion of the endogenous differentially expressed or pathway gene product would.

The differentially expressed or pathway gene products may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the differentially expressed or pathway gene polypeptides and peptides of the invention by expressing nucleic acid encoding differentially expressed or pathway gene sequences are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing differentially expressed or pathway gene protein coding sequences and appropriate transcriptional/translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques and in vivo

recombination/genetic recombination. See, for example, the techniques described in Sambrook et al., 1989, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y. which is incorporated by reference herein in their entirety, and Ausubel, 1989, supra. Alternatively, RNA capable of encoding differentially expressed or pathway gene protein sequences may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide Synthesis", 1984, Gait, M.J. ed., IRL Press, Oxford, which is incorporated by reference herein in its entirety.

A variety of host-expression vector systems may be utilized to express the differentially expressed or pathway gene coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the differentially expressed or pathway gene protein of the invention in situ. These include but are not limited to microorganisms such as bacteria (e.g., E. coli, B. subtilis) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing differentially expressed or pathway gene protein coding sequences; yeast (e.g., Saccharomyces, Pichia) transformed with recombinant yeast expression vectors containing the differentially expressed or pathway gene protein coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing the differentially expressed or pathway gene protein coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g.,

cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing differentially expressed or pathway gene protein coding sequences; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the differentially expressed or pathway gene protein being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of antibodies or to screen peptide libraries, for example, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the differentially expressed or pathway gene protein coding sequence may be ligated individually into the vector in frame with the lacZ coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. Chem. 264:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are

designed to include thrombin or factor Xa protease cleavage sites so that the cloned Target gene protein can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The differentially expressed or pathway gene coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of differentially expressed or pathway gene coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed. (E.g., see Smith et al., 1983, J. Virol. 46:584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the differentially expressed or pathway gene coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing differentially expressed or pathway gene protein in infected hosts. (e.g., See Logan & Shenk, 1984, Proc.

Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted differentially expressed or pathway gene coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire differentially expressed or pathway gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the differentially expressed or pathway gene coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., 1987, Methods in Enzymol. 153:516-544).

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of

the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, etc.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express the differentially expressed or pathway gene protein may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the differentially expressed or pathway gene protein. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the differentially expressed or pathway gene protein.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase

(Szybalska & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, et al., 1980, Cell 22:817) genes can be employed in tk⁻, hgp^{rt}- or ap^{rt}- cells, respectively.

5 Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance

10 to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., 1981, J. Mol. Biol. 150:1); and hyg^{ro}, which confers resistance to hygromycin

15 (Santerre, et al., 1984, Gene 30:147) genes.

When used as a component in assay systems such as those described herein, the differentially expressed or pathway gene protein may be labeled, either directly or indirectly, to facilitate detection of a

20 complex formed between the differentially expressed or pathway gene protein and a test substance. Any of a variety of suitable labeling systems may be used including but not limited to radioisotopes such as ¹²⁵I; enzyme labelling systems that generate a

25 detectable colorimetric signal or light when exposed to substrate; and fluorescent labels.

Where recombinant DNA technology is used to produce the differentially expressed or pathway gene protein for such assay systems, it may be advantageous

30 to engineer fusion proteins that can facilitate labeling, immobilization and/or detection.

Fusion proteins which facilitate solubility can include, but are not limited to soluble Ig-tailed fusion proteins. Methods for engineering such soluble

35 Ig-tailed fusion proteins are well known to those of

skill in the art. See, for example, U.S. Patent No. 5,116,964, which is incorporated herein by reference in its entirety.

Indirect labeling involves the use of a protein,
5 such as a labeled antibody, which specifically binds to a differentially expressed or pathway gene product. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by an Fab expression
10 library.

5.2.3. ANTIBODIES SPECIFIC FOR DIFFERENTIALLY EXPRESSED OR PATHWAY GENE PRODUCTS

Described herein are methods for the production
15 of antibodies capable of specifically recognizing one or more differentially expressed or pathway gene epitopes. Such antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies,
20 single chain antibodies, Fab fragments, F(ab')₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection
25 of a fingerprint, target, or pathway gene in a biological sample, or, alternatively, as a method for the inhibition of abnormal target gene activity. Thus, such antibodies may be utilized as part of body weight disorder treatment methods, and/or may be used
30 as part of diagnostic techniques whereby patients may be tested for abnormal levels of fingerprint, target, or pathway gene proteins, or for the presence of abnormal forms of the such proteins.

For the production of antibodies to a
35 differentially expressed or pathway gene, various host

animals may be immunized by injection with a differentially expressed or pathway gene protein, or a portion thereof. Such host animals may include but are not limited to rabbits, mice, and rats, to name
5 but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as
10 lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

15 Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as target gene product, or an antigenic functional derivative thereof. For the production of polyclonal
20 antibodies, host animals such as those described above, may be immunized by injection with differentially expressed or pathway gene product supplemented with adjuvants as also described above.

Monoclonal antibodies, which are homogeneous
25 populations of antibodies to a particular antigen, may be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to the hybridoma technique of Kohler and Milstein,
30 (1975, Nature 256:495-497; and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor et al., 1983, Immunology Today 4:72; Cole et al., 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985,
35 Monoclonal Antibodies And Cancer Therapy, Alan R.

Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci., 81:6851-6855; Neuberger et al., 1984, Nature, 312:604-606; Takeda et al., 1985, Nature, 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, Science 242:423-426; Huston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 334:544-546) can be adapted to produce differentially expressed or pathway gene-single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab'), fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the

disulfide bridges of the F(ab'), fragments.

Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

5.2.4. CELL- AND ANIMAL-BASED MODEL SYSTEMS

Described herein are cell- and animal-based systems which act as models for body weight disorders.

10 These systems may be used in a variety of applications. For example, the animal-based model systems can be utilized to identify differentially expressed genes via one of the paradigms described, above, in Section 5.1.1.1. Cell- and animal-based
15 model systems may be used to further characterize differentially expressed and pathway genes, as described, above, in Section 5.1.3. Such further characterization may, for example, indicate that a differentially expressed gene is a target gene.
20 Second, such assays may be utilized as part of screening strategies designed to identify compounds which are capable of ameliorating body weight disorder symptoms, as described, below. Thus, the animal- and cell-based models may be used to identify drugs,
25 pharmaceuticals, therapies and interventions which may be effective in treating such body weight disorders. In addition, as described in detail, below, in Section 5.6, such animal models may be used to determine the LD₅₀ and the ED₅₀ in animal subjects, and such data can
30 be used to determine the in vivo efficacy of potential body weight disorder treatments.

35

5.2.4.1. ANIMAL-BASED SYSTEMS

Animal-based model systems of body weight disorders may include, but are not limited to, non-recombinant and engineered transgenic animals.

5 Non-recombinant animal models for body weight disorders may include, for example, genetic models. Such genetic body disorder models may include, for example, mouse models of obesity such as mice homozygous for the autosomal recessive ob, db, or tub
10 alleles.

Non-recombinant, non-genetic animal models of body weight disorders may include, for example, rat models in which bilateral lesions exist in the ventromedial hypothalamus, leading to hyperphagia and
15 gross obesity, or in which ventrolateral hypothalamus lesions exist, which lead to aphagia. Further, mice which, as newborns, are fed mono-sodium-glutamate (MSG) develop obesity, and may, therefore, also be utilized as animal models for body weight disorders.

20 Additionally, animal models exhibiting body weight disorder-like symptoms may be engineered by utilizing, for example, target gene sequences such as those described, above, in Section 5.2, in conjunction with techniques for producing transgenic animals that
25 are well known to those of skill in the art. For example, target gene sequences may be introduced into, and overexpressed in, the genome of the animal of interest, or, if endogenous target gene sequences are present, they may, either be overexpressed or,
30 alternatively, may be disrupted in order to underexpress or inactivate target gene expression.

In order to overexpress a target gene sequence, the coding portion of the target gene sequence may be ligated to a regulatory sequence which is capable of
35 driving gene expression in the animal and cell type of

interest. Such regulatory regions will be well known to those of skill in the art, and may be utilized in the absence of undue experimentation.

For underexpression of an endogenous target gene sequence, such a sequence may be isolated and engineered such that when reintroduced into the genome of the animal of interest, the endogenous target gene alleles will be inactivated. Preferably, the engineered target gene sequence is introduced via gene targeting such that the endogenous target sequence is disrupted upon integration of the engineered target gene sequence into the animal's genome. Gene targeting is discussed, below, in this Section.

Animals of any species, including, but not limited to, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, goats, and non-human primates, e.g., baboons, squirrels, monkeys, and chimpanzees may be used to generate body weight disorder animal models.

Any technique known in the art may be used to introduce a target gene transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to pronuclear microinjection (Hoppe, P.C. and Wagner, T.E., 1989, U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten et al., 1985, Proc. Natl. Acad. Sci., USA 82:6148-6152); gene targeting in embryonic stem cells (Thompson et al., 1989, Cell 56:313-321); electroporation of embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814); and sperm-mediated gene transfer (Lavitrano et al., 1989, Cell 57:717-723); etc. For a review of such techniques, see Gordon, 1989, Transgenic Animals, Intl. Rev. Cytol. 115:171-229, which is incorporated by reference herein in its entirety.

35

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals. (See, 5 for example, techniques described by Jakobovits, 1994, Curr. Biol. 4: 761-763). The transgene may be integrated as a single transgene or in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into 10 and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko, M. et al., 1992, Proc. Natl. Acad. Sci. USA 89:6232-6236). The regulatory sequences required for such a cell-type specific activation will depend upon the 15 particular cell type of interest, and will be apparent to those of skill in the art.

When it is desired that the target gene transgene be integrated into the chromosomal site of the endogenous target gene, gene targeting is preferred. 20 Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous target gene of interest are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, 25 into and disrupting the function of, the nucleotide sequence of the endogenous target gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene of interest in only that cell type, by following, for 30 example, the teaching of Gu et al. (Gu, H. et al., 1994, Science 265:103-106). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

35

Once transgenic animals have been generated, the expression of the recombinant target gene and protein may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to assay whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include but are not limited to Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and RT-PCR. Samples of target gene-expressing tissue, may also be evaluated immunocytochemically using antibodies specific for the target gene transgene gene product of interest.

The target gene transgenic animals that express target gene mRNA or target gene transgene peptide (detected immunocytochemically, using antibodies directed against target gene product epitopes) at easily detectable levels should then be further evaluated to identify those animals which display characteristic body weight disorder-like symptoms. Such symptoms may include, for example, obesity, anorexia, and an abnormal food intake. Additionally, specific cell types within the transgenic animals may be analyzed and assayed for cellular phenotypes characteristic of body weight disorders. Such cellular phenotypes may include, for example, abnormal adipocyte differentiation (e.g., abnormal preadipocyte/~~adipocyte~~ differentiation) and metabolism. Further, such cellular phenotypes may include as assessment of a particular cell types fingerprint pattern of expression and its comparison to known fingerprint expression profiles of the particular cell type in animals exhibiting body weight

disorders. Such transgenic animals serve as suitable model systems for body weight disorders.

Once target gene transgenic founder animals are produced (i.e., those animals which express target gene proteins in cells or tissues of interest, and which, preferably, exhibit symptoms of body weight disorders), they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound target gene transgenics that express the target gene transgene of interest at higher levels because of the effects of additive expression of each target gene transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the possible need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; breeding animals to different inbred genetic backgrounds so as to examine effects of modifying alleles on expression of the target gene transgene and the development of body weight disorder-like symptoms. One such approach is to cross the target gene transgenic founder animals with a wild type strain to produce an F1 generation that exhibits body weight disorder-like symptoms, such as obesity, anorexia, and abnormal food intake. The F1 generation may then be inbred in order to develop a homozygous line, if it is found that homozygous target gene transgenic animals are viable.

35

5.2.4.2. CELL-BASED ASSAYS

Cells that contain and express target gene sequences which encode target gene protein, and, further, exhibit cellular phenotypes associated with a body weight disorder of interest, may be utilized to identify compounds that exhibit an ability to ameliorate body weight disorder symptoms. Cellular phenotypes which may indicate an ability to ameliorate body weight disorders may include, for example, inhibition of adipose cell differentiation (e.g., an inhibition of differentiation of preadipocytes into adipocytes) and an inhibition of the ability of adipocytes to synthesize fat.

Further, the fingerprint pattern of gene expression of cells of interest may be analyzed and compared to the normal, non-body weight disorder fingerprint pattern. Those compounds which cause cells exhibiting body weight disorder-like cellular phenotypes to produce a fingerprint pattern more closely resembling a normal fingerprint pattern for the cell of interest may be considered candidates for further testing regarding an ability to ameliorate body weight disorder symptoms.

Cells which be utilized for such assays may, for example, include non-recombinant cell lines, such as preadipocyte cell lines such as 3T3-L1 and TA1 mouse preadipocyte cell lines, liver cell lines, such as the Hepa1-6 mouse liver cell line, and the HepG2 human liver cell line.

Further, cells which may be used for such assays may also include recombinant, transgenic cell lines. For example, the body weight disorder animal models of the invention, discussed, above, in Section 5.2.4.1, may be used to generate cell lines, containing one or more cell types involved in body weight disorders,

that can be used as cell culture models for this disorder. While primary cultures derived from the body weight disorder transgenic animals of the invention may be utilized, the generation of
5 continuous cell lines is preferred. For examples of techniques which may be used to derive a continuous cell line from the transgenic animals, see Small et al., 1985, Mol. Cell Biol. 5:642-648.

Alternatively, cells of a cell type known to be
10 involved in body weight disorders may be transfected with sequences capable of increasing or decreasing the amount of target gene expression within the cell. For example, target gene sequences may be introduced into, and overexpressed in, the genome of the cell of
15 interest, or, if endogenous target gene sequences are present, they may either be overexpressed or, alternatively, be disrupted in order to underexpress or inactivate target gene expression.

In order to overexpress a target gene sequence,
20 the coding portion of the target gene sequence may be ligated to a regulatory sequence which is capable of driving gene expression in the cell type of interest. Such regulatory regions will be well known to those of skill in the art, and may be utilized in the absence
25 of undue experimentation.

For underexpression of an endogenous target gene sequence, such a sequence may be isolated and engineered such that when reintroduced into the genome of the cell type of interest, the endogenous target
30 gene alleles will be inactivated. Preferably, the engineered target gene sequence is introduced via gene targeting such that the endogenous target sequence is disrupted upon integration of the engineered target gene sequence into the cell's genome. Gene targeting
35 is discussed, above, in Section 5.4.2.1.

Transfection of target gene sequence nucleic acid may be accomplished by utilizing standard techniques. See, for example, Ausubel, 1989, supra. Transfected cells should be evaluated for the presence of the recombinant target gene sequences, for expression and accumulation of target gene mRNA, and for the presence of recombinant target gene protein production. In instances wherein a decrease in target gene expression is desired, standard techniques may be used to demonstrate whether a decrease in endogenous target gene expression and/or in target gene product production is achieved.

5.3. SCREENING ASSAYS FOR COMPOUNDS
THAT INTERACT WITH THE TARGET
GENE PRODUCT

The following assays are designed to identify compounds that bind to target gene products, bind to other cellular proteins that interact with a target gene product, and to compounds that interfere with the interaction of the target gene product with other cellular proteins. Such compounds may include, but are not limited to, other cellular proteins. Methods for the identification of such cellular proteins are described, below, in Section 5.3.2.

Compounds may include, but are not limited to, peptides such as, for example, soluble peptides, including but not limited to, Ig-tailed fusion peptides, comprising extracellular portions of target gene product transmembrane receptors, and members of random peptide libraries; (see, e.g., Lam, K.S. et al., 1991, Nature 354:82-84; Houghten, R. et al., 1991, Nature 354:84-86), made of D- and/or L-configuration amino acids, phosphopeptides (including, but not limited to members of random or partially degenerate, directed phosphopeptide libraries; see,

e.g., Songyang, Z. et al., 1993, Cell 72:767-778), antibodies (including, but not limited to, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, F(ab'), and FAb expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

Compounds identified via assays such as those described herein may be useful, for example, in elaborating the biological function of the target gene product, and for ameliorating body weight disorders. In instances, for example, whereby a body weight disorder situation results from a lower overall level of target gene expression, target gene product, and/or target gene product activity in a cell or tissue involved in such a body weight disorder, compounds that interact with the target gene product may include ones which accentuate or amplify the activity of the bound target gene protein. Such compounds would bring about an effective increase in the level of target gene activity, thus ameliorating symptoms. In instances whereby mutations within the target gene cause aberrant target gene proteins to be made which have a deleterious effect that leads to a body weight disorder, compounds that bind target gene protein may be identified that inhibit the activity of the bound target gene protein. Assays for testing the effectiveness of compounds, identified by, for example, techniques such as those described in Section 5.3.1-5.3.3, are discussed, below, in Section 5.3.4.

5.3.1. IN VITRO SCREENING ASSAYS FOR COMPOUNDS THAT BIND TO THE TARGET GENE PRODUCT

In vitro systems may be designed to identify compounds capable of binding the target gene products

of the invention. Compounds identified may be useful, for example, in modulating the activity of wild type and/or mutant target gene products, may be useful in elaborating the biological function of the target gene product, may be utilized in screens for identifying compounds that disrupt normal target gene product interactions, or may in themselves disrupt such interactions.

The principle of the assays used to identify compounds that bind to the target gene product involves preparing a reaction mixture of the target gene product and the test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex which can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways. For example, one method to conduct such an assay would involve anchoring target gene product or the test substance onto a solid phase and detecting target gene product/test compound complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, the target gene product may be anchored onto a solid surface, and the test compound, which is not anchored, may be labeled, either directly or indirectly.

In practice, microtiter plates may conveniently be utilized as the solid phase. The anchored component may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished by simply coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein to be immobilized may be used to anchor the protein to the solid

35

surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the nonimmobilized component is added to the coated surface containing
5 the anchored component. After the reaction is complete, unreacted components are removed (e.g., by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid
10 surface can be accomplished in a number of ways. Where the previously nonimmobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously nonimmobilized component is not pre-
15 labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the previously nonimmobilized component (the antibody, in turn, may be directly labeled or indirectly labeled with a
20 labeled anti-Ig antibody).

Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for target gene
25 product or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

30 5.3.2. ASSAYS FOR CELLULAR PROTEINS THAT INTERACT WITH THE TARGET GENE PROTEIN

Any method suitable for detecting protein-protein interactions may be employed for identifying novel
35 target protein-cellular or extracellular protein interactions. These methods are outlined in section

5.1.2., above, for the identification of pathway genes, and may be utilized herein with respect to the identification of proteins which interact with identified target proteins.

5

10

15

20

25

30

35

5.3.3. ASSAYS FOR COMPOUNDS THAT INTERFERE
WITH TARGET GENE PRODUCT/CELLULAR
MACROMOLECULE INTERACTION

The target gene products of the invention may, in vivo, interact with one or more cellular or extracellular macromolecules, such as proteins. Such macromolecules may include, but are not limited to, nucleic acid molecules and those proteins identified via methods such as those described, above, in Section 5.3.2. For purposes of this discussion, such cellular and extracellular macromolecules are referred to herein as "binding partners". Compounds that disrupt such interactions may be useful in regulating the activity of the target gene product, especially mutant target gene products. Such compounds may include, but are not limited to molecules such as antibodies, peptides, and the like, as described, for example, in Section 5.3.1. above.

The basic principle of the assay systems used to identify compounds that interfere with the interaction between the target gene product and its cellular or extracellular binding partner or partners involves preparing a reaction mixture containing the target gene product, and the binding partner under conditions and for a time sufficient to allow the two to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound may be initially included in the reaction mixture, or may be added at a time subsequent to the addition of target gene product and its cellular or extracellular binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the target gene protein and the cellular or extracellular binding partner is then

detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the target gene protein and the interactive binding partner.

5 Additionally, complex formation within reaction mixtures containing the test compound and normal target gene protein may also be compared to complex formation within reaction mixtures containing the test

10 compound and a mutant target gene protein. This comparison may be important in those cases wherein it is desirable to identify compounds that disrupt interactions of mutant but not normal target gene proteins.

15 The assay for compounds that interfere with the interaction of the target gene products and binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the target gene product or the

20 binding partner onto a solid phase and detecting complexes anchored on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be

25 varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the target gene products and the binding partners, e.g., by competition, can be identified by conducting the

30 reaction in the presence of the test substance; i.e., by adding the test substance to the reaction mixture prior to or simultaneously with the target gene protein and interactive cellular or extracellular binding partner. Alternatively, test compounds that

35 disrupt preformed complexes, e.g. compounds with

higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are described briefly below.

In a heterogeneous assay system, either the target gene protein or the interactive cellular or extracellular binding partner, is anchored onto a solid surface, while the non-anchored species is labeled, either directly or indirectly. In practice, microtiter plates are conveniently utilized. The anchored species may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface with a solution of the target gene product or binding partner and drying. Alternatively, an immobilized antibody specific for the species to be anchored may be used to anchor the species to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (e.g., by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. ~~Where the non-immobilized~~ species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, may be directly labeled or indirectly labeled

with a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds which inhibit complex formation or which disrupt preformed complexes can be detected.

5 Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for one of the
10 binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds which inhibit complex
15 or which disrupt preformed complexes can be identified.

 In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of the target gene protein and the
20 interactive cellular or extracellular binding partner is prepared in which either the target gene product or its binding partners is labeled, but the signal generated by the label is quenched due to complex formation (see, e.g., U.S. Patent No. 4,109,496 by
25 Rubenstein which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test
30 ~~substances which disrupt~~ target gene protein/cellular or extracellular binding partner interaction can be identified.

 In a particular embodiment, the target gene product can be prepared for immobilization using
35 recombinant DNA techniques described in Section 5.2.1,

above. For example, the target gene coding region can be fused to a glutathione-S-transferase (GST) gene using a fusion vector, such as pGEX-5X-1, in such a manner that its binding activity is maintained in the resulting fusion protein. The interactive cellular or extracellular binding partner can be purified and used to raise a monoclonal antibody, using methods routinely practiced in the art and described above, in Section 5.2.3. This antibody can be labeled with the radioactive isotope ^{125}I , for example, by methods routinely practiced in the art. In a heterogeneous assay, e.g., the GST-target gene fusion protein can be anchored to glutathione-agarose beads. The interactive cellular or extracellular binding partner can then be added in the presence or absence of the test compound in a manner that allows interaction and binding to occur. At the end of the reaction period, unbound material can be washed away, and the labeled monoclonal antibody can be added to the system and allowed to bind to the complexed components. The interaction between the target gene protein and the interactive cellular or extracellular binding partner can be detected by measuring the amount of radioactivity that remains associated with the glutathione-agarose beads. A successful inhibition of the interaction by the test compound will result in a decrease in measured radioactivity.

Alternatively, the GST-target gene fusion protein and the interactive cellular or extracellular binding partner can be mixed together in liquid in the absence of the solid glutathione-agarose beads. The test compound can be added either during or after the species are allowed to interact. This mixture can then be added to the glutathione-agarose beads and unbound material is washed away. Again the extent of

inhibition of the target gene product/binding partner interaction can be detected by adding the labeled antibody and measuring the radioactivity associated with the beads.

5 In another embodiment of the invention, these same techniques can be employed using peptide fragments that correspond to the binding domains of the target gene protein and/or the interactive cellular or extracellular binding partner (in cases
10 where the binding partner is a protein), in place of one or both of the full length proteins. Any number of methods routinely practiced in the art can be used to identify and isolate the binding sites. These methods include, but are not limited to, mutagenesis
15 of the gene encoding one of the proteins and screening for disruption of binding in a co-immunoprecipitation assay. Compensating mutations in the gene encoding the second species in the complex can then be selected. Sequence analysis of the genes encoding the
20 respective proteins will reveal the mutations that correspond to the region of the protein involved in interactive binding. Alternatively, one protein can be anchored to a solid surface using methods described in this Section above, and allowed to interact with
25 and bind to its labeled binding partner, which has been treated with a proteolytic enzyme, such as trypsin. After washing, a short, labeled peptide comprising the binding domain may remain associated with the solid material, which can be isolated and
30 identified by amino acid sequencing. Also, once the gene coding for the for the cellular or extracellular binding partner is obtained, short gene segments can be engineered to express peptide fragments of the protein, which can then be tested for binding activity
35 and purified or synthesized.

For example, and not by way of limitation, a target gene product can be anchored to a solid material as described, above, in this Section by making a GST-target gene fusion protein and allowing it to bind to glutathione agarose beads. The interactive cellular or extracellular binding partner can be labeled with a radioactive isotope, such as ^{35}S , and cleaved with a proteolytic enzyme such as trypsin. Cleavage products can then be added to the anchored GST-target gene fusion protein and allowed to bind. After washing away unbound peptides, labeled bound material, representing the cellular or extracellular binding partner binding domain, can be eluted, purified, and analyzed for amino acid sequence by well-known methods. Peptides so identified can be produced synthetically or fused to appropriate facilitative proteins using recombinant DNA technology.

5.3.4. ASSAYS FOR AMELIORATION OF BODY WEIGHT DISORDER SYMPTOMS

Any of the binding compounds, including but not limited to, compounds such as those identified in the foregoing assay systems, may be tested for the ability to ameliorate body weight disorder symptoms, which may include, for example, obesity, anorexia, and/or an abnormal level of food intake. Cell-based and animal model-based assays for the identification of compounds exhibiting such an ability to ameliorate body weight disorder symptoms are described below.

First, cell-based systems such as those described, above, in Section 5.2.4.2, may be used to identify compounds which may act to ameliorate body weight disorder symptoms. For example, such cell systems may be exposed to a compound, suspected of exhibiting an ability to ameliorate body weight

disorder symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of body weight disorder symptoms in the exposed cells. After exposure, the cells are examined to determine
5 whether one or more of the body weight disorder-like cellular phenotypes has been altered to resemble a more normal or more wild type, non-body weight disorder phenotype, or a phenotype more likely to produce a lower incidence or severity of disorder
10 symptoms.

In addition, animal-based body weight disorder systems, such as those described, above, in Section 5.2.4.1, may be used to identify compounds capable of ameliorating body weight disorder-like symptoms. Such
15 animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies and interventions which may be effective in treating such disorders. For example, animal models may be exposed to a compound, suspected of exhibiting an
20 ability to ameliorate body weight disorder symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of body weight disorder symptoms in the exposed animals. The response of the animals to the exposure may be
25 monitored by assessing the reversal of disorders associated with body weight disorders such as obesity.

With regard to intervention, any treatments which reverse any aspect of body weight disorder-like symptoms should be considered as candidates for human
30 body weight disorder therapeutic intervention. Dosages of test agents may be determined by deriving dose-response curves, as discussed in Section 5.6.1, below.

Gene expression patterns may be utilized in
35 conjunction with either cell-based or animal-based

systems to assess the ability of a compound to ameliorate body weight disorder-like symptoms. For example, the expression pattern of one or more fingerprint genes may form part of a fingerprint profile which may be then be used in such an assessment. Fingerprint profiles are described, below, in Section 5.7.1. Fingerprint profiles may be characterized for known states, either body weight disorder or normal states, within the cell- and/or animal-based model systems. Subsequently, these known fingerprint profiles may be compared to ascertain the effect a test compound has to modify such fingerprint profiles, and to cause the profile to more closely resemble that of a more desirable fingerprint. For example, administration of a compound may cause the fingerprint profile of a body weight disorder model system to more closely resemble the control system. Administration of a compound may, alternatively, cause the fingerprint profile of a control system to begin to mimic a body weight disorder state, which may, for example, be used in further characterizing the compound of interest, or may be used in the generation of additional animal models.

25 5.4. COMPOUNDS AND METHODS FOR TREATMENT OF BODY WEIGHT DISORDERS

Described below are methods and compositions whereby body weight disorder symptoms may be ameliorated. It is possible that body weight disorders may be brought about, at least in part, by an abnormal level of target gene product, or by the presence of a target gene product exhibiting an abnormal activity. As such, the reduction in the level and/or activity of such target gene products would bring about the amelioration of body weight disorder-like symptoms. Techniques for the reduction

of target gene expression levels or target gene product activity levels are discussed in Section 5.4.1, below.

Alternatively, it is possible that body weight disorders may be brought about, at least in part, by the absence or reduction of the level of target gene expression, or a reduction in the level of a target gene product's activity. As such, an increase in the level of target gene expression and/or the activity of such gene products would bring about the amelioration of body weight disorder-like symptoms. Techniques for increasing target gene expression levels or target gene product activity levels are discussed in Section 5.4.2, below.

15

5.4.1. COMPOUNDS THAT INHIBIT EXPRESSION,
SYNTHESIS OR ACTIVITY OF MUTANT
TARGET GENE ACTIVITY

As discussed above, target genes involved in body weight disorders may cause such disorders via an increased level of target gene activity. A variety of techniques may be utilized to inhibit the expression, synthesis, or activity of such target genes and/or proteins.

For example, compounds such as those identified through assays described, above, in Section 5.3, which exhibit inhibitory activity, may be used in accordance with the invention to ameliorate body weight disorder symptoms. As discussed in Section 5.3, above, such molecules may include, but are not limited to, peptides, (such as, for example, peptides representing soluble extracellular portions of target gene product transmembrane receptors), phosphopeptides, small organic or inorganic molecules, or antibodies (including, for example, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain

antibodies, and FAb, F(ab'), and FAb expression library fragments, and epitope-binding fragments thereof).

Techniques for determination of effective doses and administration of such compounds are described below, in Section 5.6.1. Inhibitory antibody techniques are further described, below, in Section 5.4.1.2.

Further, antisense and ribozyme molecules which inhibit expression of the target gene may also be used in accordance with the invention to inhibit the aberrant target gene activity. Such techniques are described, below, in Section 5.4.1.1. Still further, as described, below, in Section 5.4.1.1, triple helix molecules may be utilized in inhibiting the aberrant target gene activity.

5.4.1.1. INHIBITORY ANTISENSE, RIBOZYME AND TRIPLE HELIX APPROACHES

Among the compounds which may exhibit the ability to ameliorate body weight disorder symptoms are antisense, ribozyme, and triple helix molecules. Such molecules may be designed to reduce or inhibit either wild type, or if appropriate, mutant target gene activity. Techniques for the production and use of such molecules are well known to those of skill in the art.

Anti-sense RNA and DNA molecules act to directly block the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. With respect to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, e.g., between the -10 and +10 regions of the target gene nucleotide sequence of interest, are preferred.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. (For a review, see Rossi, J., 1994, Current Biology 4:469-

- 471). The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage. The composition of ribozyme molecules must include one or more sequences complementary to the target gene mRNA, and must include the well known catalytic sequence responsible for mRNA cleavage. For this sequence, see U.S. Pat. No. 5,093,246, which is incorporated by reference herein in its entirety. As such, within the scope of the invention are engineered hammerhead motif ribozyme molecules that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences encoding target gene proteins.
- Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the molecule of interest for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for predicted structural features, such as secondary structure, that may render the oligonucleotide sequence unsuitable. The suitability of candidate sequences may also be evaluated by testing their accessibility to hybridize with complementary oligonucleotides, using ribonuclease protection assays.
- ~~Nucleic acid molecules to be used in triplex~~ helix formation for the inhibition of transcription should be single stranded and composed of deoxynucleotides. The base composition of these oligonucleotides must be designed to promote triple helix formation via Hoogsteen base pairing rules,

which generally require sizeable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC* triplets

5 across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules

10 may be chosen that are purine-rich, for example, contain a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex,

15 resulting in GGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic

20 acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present

25 on one strand of a duplex.

In instances wherein the antisense, ribozyme, and/or triple helix molecules described herein are utilized to inhibit mutant gene expression, it is possible that the technique may so efficiently reduce

30 or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles that the possibility may arise wherein the concentration of normal target gene product present may be lower than is necessary for a

35 normal phenotype. In such cases, to ensure that

substantially normal levels of target gene activity are maintained, therefore, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity may, be introduced into
5 cells via gene therapy methods such as those described, below, in Section 5.5. that do not contain sequences susceptible to whatever antisense, ribozyme, or triple helix treatments are being utilized. Alternatively, in instances whereby the target gene
10 encodes an extracellular protein, it may be preferable to coadminister normal target gene protein in order to maintain the requisite level of target gene activity.

Anti-sense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by
15 any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical
20 synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase
25 promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

30 Various well-known modifications to the DNA molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to, the addition of flanking sequences of ribo- or deoxy-
35 nucleotides to the 5' and/or 3' ends of the molecule

5 5.4.1.2. ANTIBODIES FOR INHIBITION OF
 TARGET GENE PRODUCTS

20 In instances where the target gene protein is
intracellular and whole antibodies are used,
internalizing antibodies may be preferred. However,
lipofectin or liposomes may be used to deliver the
antibody or a fragment of the Fab region which binds
25 to the target gene product epitope into cells. Where
fragments of the antibody are used, the smallest
inhibitory fragment which binds to the target
protein's binding domain is preferred. For example,
peptides having an amino acid sequence corresponding
30 to the domain of the variable region of the antibody
that binds to the target gene protein may be used.
Such peptides may be synthesized chemically or
produced via recombinant DNA technology using methods
well known in the art (e.g., see Creighton, 1983,
35 supra; and Sambrook et al., 1989, supra).

Alternatively, single chain neutralizing antibodies which bind to intracellular target gene product epitopes may also be administered. Such single chain antibodies may be administered, for
5 example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population by utilizing, for example, techniques such as those described in Marasco et al. (Marasco, W. et al., 1993, Proc. Natl. Acad. Sci. USA 90:7889-7893).
10 In instances where the target gene protein is extracellular, or is a transmembrane protein, any of the administration techniques described, below in Section 5.6 which are appropriate for peptide administration may be utilized to effectively
15 administer inhibitory target gene antibodies to their site of action.

5.4.2. METHODS FOR RESTORING TARGET GENE ACTIVITY

Target genes that cause body weight disorders may
20 be underexpressed within body weight disorder situations. Alternatively, the activity of target gene products may be diminished, leading to the development of body weight disorder symptoms. Described in this Section are methods whereby the
25 level of target gene activity may be increased to levels wherein body weight disorder symptoms are ameliorated. The level of gene activity may be increased, for example, by either increasing the level of target gene product present or by increasing the
30 level of active target gene product which is present.

For example, a target gene protein, at a level sufficient to ameliorate body weight disorder symptoms may be administered to a patient exhibiting such symptoms. Any of the techniques discussed, below, in
35 Section 5.6, may be utilized for such administration.

One of skill in the art will readily know how to determine the concentration of effective, non-toxic doses of the normal target gene protein, utilizing techniques such as those described, below, in Section 5.6.1.

Further, patients may be treated by gene replacement therapy. One or more copies of a normal target gene or a portion of the gene that directs the production of a normal target gene protein with target gene function, may be inserted into cells, using vectors which include, but are not limited to adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles that introduce DNA into cells, such as liposomes. Additionally, techniques such as those described above may be utilized for the introduction of normal target gene sequences into human cells.

Cells, preferably, autologous cells, containing normal target gene expressing gene sequences may then be introduced or reintroduced into the patient at positions which allow for the amelioration of body weight disorder symptoms. Such cell replacement techniques may be preferred, for example, when the target gene product is a secreted, extracellular gene product.

Additionally, antibodies may be administered which specifically bind to a target protein and, by binding, serve to, either directly or indirectly, activate the target protein function. Such antibodies can include, but are not limited to polyclonal, monoclonal, FAb fragments, single chain antibodies, chimeric antibodies and the like. The antibodies may be generated using standard techniques such as those described, above, in Section 5.2.3., and may be generated against the protein themselves or against

proteins corresponding to portions of the proteins.
The antibodies may be administered, for example,
according to the techniques described, above, in
Section 5.4.1.2.

5

5.6. PHARMACEUTICAL PREPARATIONS AND METHODS OF ADMINISTRATION

The identified compounds, nucleic acid molecules
and cells that affect target gene expression,
10 synthesis and/or activity can be administered to a
patient at therapeutically effective doses to treat or
ameliorate body weight disorders. A therapeutically
effective dose refers to that amount of the compound
sufficient to result in amelioration of symptoms of
15 body weight disorder, or alternatively, to that amount
of a nucleic acid molecule sufficient to express a
concentration of gene product which results in the
amelioration of such symptoms.

20

5.6.1. EFFECTIVE DOSE

Toxicity and therapeutic efficacy of such
compounds can be determined by standard pharmaceutical
procedures in cell cultures or experimental animals,
e.g., for determining the LD₅₀ (the dose lethal to 50%
25 of the population) and the ED₅₀ (the dose
therapeutically effective in 50% of the population).
The dose ratio between toxic and therapeutic effects
is the therapeutic index and it can be expressed as
the ratio LD₅₀/ED₅₀. Compounds which exhibit large
30 therapeutic indices are preferred. While compounds
that exhibit toxic side effects may be used, care
should be taken to design a delivery system that
targets such compounds to the site of affected tissue
in order to minimize potential damage to uninfected
35 cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of
5 circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any
10 compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound which
15 achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

20

5.6.2. FORMULATIONS AND USE

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically
25 acceptable carriers or excipients.

Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or oral, buccal,
30 parenteral or rectal administration.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as
35 binding agents (*e.g.*, pregelatinised maize starch,

polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycollate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane,

dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g. gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

5.7. DIAGNOSIS OF BODY WEIGHT DISORDER ABNORMALITIES

A variety of methods may be employed for the diagnosis of body weight disorders, predisposition to body weight disorders, for monitoring the efficacy of antibody weight disorder compounds during, for example, clinical trials and for monitoring patients undergoing clinical evaluation for the treatment of such body weight disorders.

Such methods may, for example, utilize reagents such as the fingerprint gene nucleotide sequences described in Sections 5.1, and antibodies directed against differentially expressed and pathway gene peptides, as described, above, in Sections 5.1.3 (peptides) and 5.2.3 (antibodies). Specifically, such reagents may be used, for example, for: (1) the detection of the presence of target gene mutations, or the detection of either over- or under-expression of target gene mRNA relative to the non-body weight disorder state; and (2) the detection of either an over- or an under-abundance of target gene product relative to the non-body weight disorder state.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific fingerprint gene nucleic acid or anti-fingerprint gene antibody reagent described herein, which may be conveniently used,

e.g., in clinical settings, to diagnose patients exhibiting body weight disorder abnormalities.

Any cell type or tissue in which the fingerprint gene is expressed may be utilized in the diagnostics
5 described below.

Among the methods which can be utilized herein are methods for monitoring the efficacy of compounds in clinical trails for the treatment of body weight disorders. Such compounds can, for example, be
10 compounds such as those described, above, in Section 5.4. Such a method comprises detecting, in a patient sample, a gene transcript or gene product which is differentially expressed in a body weight disorder state relative to its expression in a normal, or non-
15 body weight disorder, state.

Any of the nucleic acid detection techniques described, below, in Section 5.7.1 or any of the peptide detection described, below, in Section 5.7.2 can be used to detect the gene transcript or gene
20 product which is differentially expressed in a body weight disorder relative to its expression in the normal, or non-immune disorder, state.

During clinical trials, for example, the expression of a single fingerprint gene, or
25 alternatively, a fingerprint pattern of a cell involved in a body weight disorder can be determined in the presence or absence of the compound being tested. The efficacy of the compound can be followed by comparing the expression data obtained to the
30 corresponding known ~~expression patterns in a normal,~~ non-body weight disorder state. Compounds exhibiting efficacy are those which alter the single fingerprint gene expression and/or the fingerprint pattern to more closely resemble that of the normal, non-body weight
35 disorder state.

The detection of the product or products of genes differentially expressed in a body weight disorder state relative to their expression in a normal, or non-body weight disorder, state can also be used for monitoring the efficacy of potential anti-body weight disorder compounds during clinical trials. During clinical trials, for example, the level and/or activity of the products of one or more such differentially expressed genes can be determined in relevant cells and/or tissues in the presence or absence of the compound being tested. The efficacy of the compound can be followed by comparing the protein level and/or activity data obtained to the corresponding known levels/activities for the cells and/or tissues in a normal, non-body weight disorder state. Compounds exhibiting efficacy are those which alter the pattern of the cell and/or tissue involved in the body weight disorder to more closely resemble that of the normal, non-body weight disorder state.

20

5.7.1. DETECTION OF FINGERPRINT GENE
NUCLEIC ACIDS

DNA or RNA from the cell type or tissue to be analyzed may easily be isolated using procedures which are well known to those in the art. Diagnostic procedures may also be performed "in situ" directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents such as those described in Section 5.1 may be used as probes and/or primers for such in situ procedures (see, for example, Nuovo, G.J., 1992, "PCR In Situ Hybridization: Protocols And Applications", Raven Press, NY).

Fingerprint gene nucleotide sequences, either RNA or DNA, may, for example, be used in hybridization or

amplification assays of biological samples to detect body weight disorder-related gene structures and expression. Such assays may include, but are not limited to, Southern or Northern analyses, single
5 stranded conformational polymorphism analyses, in situ hybridization assays, and polymerase chain reaction analyses. Such analyses may reveal both quantitative aspects of the expression pattern of the fingerprint gene, and qualitative aspects of the fingerprint gene
10 expression and/or gene composition. That is, such techniques may include, for example, point mutations, insertions, deletions, chromosomal rearrangements, and/or activation or inactivation of gene expression.

Diagnostic methods for the detection of
15 fingerprint gene-specific nucleic acid molecules may involve for example, contacting and incubating nucleic acids, derived from the cell type or tissue being analyzed, with one or more labeled nucleic acid reagents as are described in Section 5.1, under
20 conditions favorable for the specific annealing of these reagents to their complementary sequences within the nucleic acid molecule of interest. Preferably, the lengths of these nucleic acid reagents are at least 15 to 30 nucleotides. After incubation, all
25 non-annealed nucleic acids are removed from the nucleic acid:fingerprint molecule hybrid. The presence of nucleic acids from the cell type or tissue which have hybridized, if any such molecules exist, is then detected. Using such a detection scheme, the
30 nucleic acid from the cell type or tissue of interest may be immobilized, for example, to a solid support such as a membrane, or a plastic surface such as that on a microtiter plate or polystyrene beads. In this case, after incubation, non-annealed, labeled nucleic
35 acid reagents of the type described in Section 5.1 are

easily removed. Detection of the remaining, annealed, labeled fingerprint nucleic acid reagents is accomplished using standard techniques well-known to those in the art.

5 Alternative diagnostic methods for the detection of fingerprint gene specific nucleic acid molecules may involve their amplification, e.g., by PCR (the experimental embodiment set forth in Mullis, K.B., 1987, U.S. Patent No. 4,683,202), ligase chain
10 reaction (Barany, F., 1991, Proc. Natl. Acad. Sci. USA 88:189-193), self sustained sequence replication (Guatelli, J.C. et al., 1990, Proc. Natl. Acad. Sci. USA 87:1874-1878), transcriptional amplification
15 system (Kwoh, D.Y. et al., 1989, Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. et al., 1988, Bio/Technology 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These
20 detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

 In one embodiment of such a detection scheme, a
25 cDNA molecule is obtained from an RNA molecule of interest (e.g., by reverse transcription of the RNA molecule into cDNA). Cell types or tissues from which such RNA may be isolated include any tissue in which wild type fingerprint gene is known to be expressed. A sequence within the cDNA is then used as the
30 template for a nucleic acid amplification reaction, such as a PCR amplification reaction, or the like. The nucleic acid reagents used as synthesis initiation reagents (e.g., primers) in the reverse transcription and nucleic acid amplification steps of this method
35 are chosen from among the fingerprint gene nucleic

acid reagents described in Section 5.1. The preferred lengths of such nucleic acid reagents are at least 9-30 nucleotides. For detection of the amplified product, the nucleic acid amplification may be performed using radioactively or non-radioactively labeled nucleotides. Alternatively, enough amplified product may be made such that the product may be visualized by standard ethidium bromide staining or by utilizing any other suitable nucleic acid staining method.

In addition to methods which focus primarily on the detection of one fingerprint nucleic acid sequence, fingerprint patterns or profiles may also be assessed in such detection schemes. "Fingerprint pattern" or "fingerprint profile", as used herein, refers to the pattern of mRNA expression obtained for a given tissue or cell type under a given set of conditions, and includes the mRNA expression of at least two of the genes within the tissue or cell type. Such conditions may include, but are not limited to body weight disorders, including obesity, and conditions relevant to processes involved in body weight or appetite regulation, including any of the control or experimental conditions described in the paradigms of Section 5.1.1.1, above. Fingerprint profiles may be generated, for example, by utilizing a differential display procedure, as discussed, above, in Section 5.1.1.2, Northern analysis and/or RT-PCR. Any of the gene sequences described, above, in Section 5.2.1 may be used as probes and/or PCR primers for the generation and corroboration of such fingerprint profiles.

35

5.7.2. DETECTION OF TARGET GENE PEPTIDES

Antibodies directed against wild type or mutant fingerprint gene peptides, which are discussed, above, in Section 5.2.3, may also be used as body weight disorder diagnostics and prognostics, as described, for example, herein. Such diagnostic methods, may be used to detect abnormalities in the level of fingerprint gene protein expression, or abnormalities in the structure and/or temporal, tissue, cellular, or subcellular location of fingerprint gene protein. Structural differences may include, for example, differences in the size, electronegativity, or antigenicity of the mutant fingerprint gene protein relative to the normal fingerprint gene protein.

Protein from the tissue or cell type to be analyzed may easily be isolated using techniques which are well known to those of skill in the art. The protein isolation methods employed herein may, for example, be such as those described in Harlow and Lane (Harlow, E. and Lane, D., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), which is incorporated herein by reference in its entirety.

Preferred diagnostic methods for the detection of wild type or mutant fingerprint gene peptide molecules may involve, for example, immunoassays wherein fingerprint gene peptides are detected by their interaction with an anti-fingerprint gene product-specific antibody.

For example, antibodies, or fragments of antibodies, such as those described, above, in Section 5.2.3, useful in the present invention may be used to quantitatively or qualitatively detect the presence of wild type or mutant fingerprint gene peptides. This can be accomplished, for example, by

immunofluorescence techniques employing a fluorescently labeled antibody (see below, this section) coupled with light microscopic, flow cytometric, or fluorimetric detection. Such techniques are especially preferred if the fingerprint gene peptides are expressed on the cell surface.

5 The antibodies (or fragments thereof) useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for in situ detection of fingerprint gene peptides. In situ detection may be accomplished by removing a labeled antibody of a patient, and applying thereto a labeled antibody of the present invention. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of the fingerprint gene peptides, but also their distribution in the examined tissue. Using the present invention, those of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such in situ detection.

25 Immunoassays for wild type or mutant fingerprint gene peptides typically comprise incubating a biological sample, such as a biological fluid, a tissue extract, freshly harvested cells, or cells which have been incubated in tissue culture, in the presence of a detectably labeled antibody capable of identifying fingerprint gene peptides, and detecting the bound antibody by any of a number of techniques well-known in the art.

30 The biological sample may be brought in contact with and immobilized onto a solid phase support or

35

carrier such as nitrocellulose, or other solid support which is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the
5 detectably labeled fingerprint gene specific antibody. The solid phase support may then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on solid support may then be detected by conventional means.

10 By "solid phase support or carrier" is intended any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified
15 celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so
20 long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the
25 surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine
30 experimentation.

The binding activity of a given lot of anti-wild type or anti-mutant fingerprint gene product antibody may be determined according to well known methods. Those skilled in the art will be able to determine

35

operative and optimal assay conditions for each determination by employing routine experimentation.

One of the ways in which the fingerprint gene peptide-specific antibody can be detectably labeled is by linking the same to an enzyme and use in an enzyme immunoassay (EIA) (Voller, A., "The Enzyme Linked Immunosorbent Assay (ELISA)", 1978, Diagnostic Horizons 2:1-7, Microbiological Associates Quarterly Publication, Walkersville, MD); Voller, A. et al., 1978, J. Clin. Pathol. 31:507-520; Butler, J.E., 1981, Meth. Enzymol. 73:482-523; Maggio, E. (ed.), 1980, ENZYME IMMUNOASSAY, CRC Press, Boca Raton, FL,; Ishikawa, E. et al., (eds.), 1981, ENZYME IMMUNOASSAY, Kagaku Shoin, Tokyo). The enzyme which is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety which can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes which can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by

radioactively labeling the antibodies or antibody fragments, it is possible to detect fingerprint gene wild type or mutant peptides through the use of a radioimmunoassay (RIA) (see, for example, Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986, which is incorporated by reference herein). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using fluorescence emitting metals such as ^{152}Eu , or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the antibody of the present invention.

Bioluminescence is a type of chemiluminescence found in biological systems in, which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is
5 determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

6. EXAMPLE: IDENTIFICATION AND
10 CHARACTERIZATION OF AN
 OBESITY-RELATED GENE

In the Example presented in this Section, one of the paradigms described, above, in Section 5.1.1.1, the genetic obesity paradigm, is utilized to identify
15 a gene which is not only differentially expressed in genetically obese test animals, but is identical to a gene which has previously been implicated in processes involved in body weight regulation. Thus, the
20 successful identification, here, of this gene corroborates the usefulness of the paradigm approach of the invention for the identification of genes involved in body weight disorders and/or in body weight or appetite regulation.

6.1. MATERIALS AND METHODS
25 Genetic obesity paradigms: 15 female C57B1/6J ob/ob mice and lean littermate controls (15 female C57B1/6J ?/+) and 15 male C57B1/Ks db/db mice and lean
30 littermate controls (15 male C57B1/ks +/+) were received from Jackson labs at 4.5 weeks of age, and housed individually on normal mouse chow (West, D.B., 1992, Am. J. Physiol. 262:R1025-R1032) for 1 week
prior to the initiation of the study. The four groups of 15 mice each were then sacrificed by CO₂ euthanasia
35 and tissues were collected. Body weight (grams) of

the four groups of mice at the time of sacrifice was measured.

- Tissue collection and RNA isolation: Following
- 5 CO₂ asphyxiation, tissues were removed and quick frozen on dry ice. Samples within an experimental or control group (15 animals per group) were then homogenized together with a mortar and pestle under liquid nitrogen.
- 10 Total cellular RNA was extracted from tissue with either RNazol™ or RNazolB™ (Tel-Test, Friendswood, TX), according to the manufacturer's instructions. Briefly, the tissue was solubilized in an appropriate amount of RNazol™ or RNazolB™, and RNA was extracted
- 15 by the addition of 1/10 v/v chloroform to the solubilized sample followed by vigorous shaking for approximately 15 seconds. The mixture was then centrifuged for 15 minutes at 12,000g and the aqueous phase was removed to a fresh tube. RNA was
- 20 precipitated with isopropanol. The resultant RNA pellet was dissolved in water and re-extracted with an equal volume of chloroform to remove any remaining phenol. The extracted volume was precipitated with 2 volumes of ethanol in the presence of 150mM sodium
- 25 acetate. The precipitated RNA was dissolved in water and the concentration determined spectroscopically (A₂₆₀).

- Differential display: Total cellular RNA (10-
- 30 50μg) was treated with 20 Units DNase I (Boehringer Mannheim, Germany) in the presence of 40 Units ribonuclease inhibitor (Boehringer Mannheim, Germany). After extraction with phenol/chloroform and ethanol precipitation, the RNA was dissolved in DEPC (diethyl
- 35 pyrocarbonate)-treated water.

Differential mRNA display was carried out as described, above, in Section 5.1.1.2. RNA (0.4-2 μ g) was reverse-transcribed using Superscript reverse transcriptase (GIBCO/BRL). The cDNAs were then
5 amplified by PCR on a Perkin-Elmer 9600 thermal cycler. The reaction mixtures (20 μ l) included arbitrary decanucleotides and one of twelve possible T₁₁VN sequences, wherein V represents either dG, dC, or dA, and N represents either dG, dT, dA, or dC.
10 Parameters for the 40 cycle PCR were as follows: Hold 94° C. 2 minutes; Cycle (40 rounds) 94° C. 15 seconds, 40° C. 2 minutes; Ramp to 72° 30 seconds; Hold 72° C. 5 minutes; Hold 4° C.

Radiolabelled PCR amplification products were
15 analyzed by electrophoresis on 6% denaturing polyacrylamide gels.

Reamplification and subcloning: PCR bands of interest were recovered from sequencing gels and
20 reamplified.

Briefly, autoradiograms were aligned with the dried gel, and the region containing the bands of interest was excised with a scalpel. The excised gel fragment was eluted by soaking in 100 μ l TE (Tris-EDTA)
25 buffer at approximately 100° C. for 15 minutes. The gel slice was then pelleted by brief centrifugation and the supernatant was transferred to a new microcentrifuge tube. DNA was combined with ethanol in the presence of 100mM Sodium acetate and 30 μ g
30 glycogen (Boehringer Mannheim, Germany) and precipitated on dry ice for approximately 10 minutes. Samples were centrifuged for 10 minutes and pellets were washed with 80% ethanol. Pellets were resuspended in 10 μ l distilled water.

35

5 μ l of the eluted DNA were reamplified in a 100 μ l reaction containing: standard Cetus Taq polymerase buffer, 20 μ M dNTPs, 1 μ M of each of the oligonucleotide primers used in the initial generation of the amplified DNA. Cycling conditions used were the same as the initial conditions used to generate the amplified band, as described above. One-half of the amplification reaction was run on a 2% agarose gel and eluted using DE-81 paper (Whatman Paper, Ltd., England) as described in Sambrook et al., supra. Recovered fragments were ligated into the cloning vector pCR[™]II (Invitrogen, Inc., San Diego CA) and transformed into competent E. coli strain DH5 α (Gibco/BRL, Gaithersburg, MD). Colonies were grown on LB-agar plates containing ampicillin (100 μ g/ml) and X-gal (40 μ g/ml) to permit blue/white selection.

Sequence analysis: After subcloning, reamplified cDNA fragments were sequenced on an Applied Biosystems Automated Sequencer (Applied Biosystems, Inc. Seattle, WA). Sequence was obtained from four independent transformants containing the same insert. The nucleotide sequence shown herein represents the consensus of the information obtained from the four sequences. Such primary sequence data was edited and trimmed of vector sequences and highly repetitive sequences and used to search Genbank databases using the BLAST (Altschul, S.F. et al., 1990, J. Mol. Biol. 215:403-410) program.

Northern analysis: RNA samples were electrophoresed in a denaturing agarose gel containing 1-1.5% agarose (SeaKem[™] LE, FMC BioProducts, Rockland, ME) containing 6.3% formaldehyde. Samples containing 5-20 μ g of total RNA were mixed with denaturing loading

solution (72% deionized formamide and bromophenol blue) and heated to 70°C. for 5 minutes. Samples were placed on ice and immediately loaded onto gels. Gels were run in 1x MOPS buffer (100mM MOPS, 25mM sodium acetate, 5mM EDTA). After electrophoresis, the gels were stained with ethidium bromide and visualized with ultraviolet light.

After completion of electrophoresis, gels were soaked in 50mM sodium hydroxide with gentle agitation for approximately 30 minutes to lightly cleave RNA. Gels were rinsed twice in water and then neutralized by soaking in 0.1M Tris-HCl (pH 7.5) for approximately 30 minutes. Gels were briefly equilibrated with 20x SSC (3M sodium chloride, 0.3M sodium citrate) and then transferred to nylon membranes such as Hybond™, -N, (Amersham, Inc., Arlington Heights, IL) or Zeta-Probe (Bio-Rad, Inc., Hercules, CA) overnight in 20x SSC. Membranes containing transferred RNA were baked at 80°C. for 2 hours to immobilize the RNA.

DNA fragments to be used as probes were of various sizes and were labeled using a random hexamer labeling technique. Briefly, 25ng of a purified DNA fragment was used to generate each probe. Fragments were added to a 20μl random hexanucleotide labeling reaction (Boehringer Mannheim, Inc., Indianapolis, IN) containing random hexamers and a mix of the nucleotides dCTP, dGTP, and dTTP (at a final concentration of 25μM each). The reaction mix was heat-denatured at 100°C. for 10 minutes and then chilled on ice. 5μl of α-³²P-dATP (50μCi; Amersham, Inc., Arlington Heights, IL) and Klenow DNA polymerase (2 units; Boehringer Mannheim, Inc., Indianapolis, IN) were added. Reactions were incubated at 37° for 30 minutes. Following incubation, 30μl water was added to the labeling reaction and unincorporated

nucleotides were removed by passing the reactions through a BioSpin-6™ chromatography column (Bio-Rad, Inc., Hercules, CA). Specific incorporation was determined using a scintillation counter. 1-5x10⁶ cpm were used per ml hybridization mixture.

Nylon membranes containing immobilized RNA were prehybridized according to manufacturer's instructions. Radiolabelled probes were heat denatured at 70°C. in 50% deionized formamide for 10 minutes and then added to the hybridization mixture (containing 50% formamide, 10% dextran sulfate, 0.1% SDS, 100µg/ml sheared salmon sperm DNA, 5x SSC, 5x Denhardt's solution, 30mM Tris-HCl (pH 8.5), 50mM NaPO₄ (pH 6.5). Hybridizations were carried out at 42°C. overnight. Nylon membranes were then bathed for 2 minutes in a wash solution of 0.2x SSC and 0.1% SDS at room temperature to remove most of the remaining hybridization solution. The membranes were then bathed twice in fresh 42°C. preheated wash solution for 20 minutes. Filters were covered in plastic wrap and exposed to autoradiographic film to visualize results.

6.2. RESULTS

Genetic obesity paradigms were utilized to identify genes which are differentially expressed in obese versus lean mice. Specifically, ob/ob and db/db obese mice were utilized in conjunction with lean littermate control mice, as described, above, in Section 6.1.

RNA samples isolated from liver tissue of the ob/ob, db/db, and littermate control mice were analyzed via differential display techniques. FIG. 1 shows amplified fragments obtained from these tissues when subjected to PCR with 11 separate primer pair combinations. The arrow in FIG. 1, indicates a PCR

product, designated band L36, which was judged to be differentially expressed among the lean and obese (ob and db) samples, with a larger amount of expression in the obese relative to the lean control samples.

5 To confirm the putative differential gene regulation, the amplified L36 band was recovered, reamplified, and used to probe Northern RNA blots which were prepared with the original liver RNA samples. FIG. 2 shows the results of one such
10 Northern blot analysis, in which the steady messages corresponding to cDNA band L36 are shown to be significantly increased in RNA samples derived from both ob/ob and db/db mice compared to lean littermate controls. Thus, this study confirmed the putative
15 differential regulation which had been suggested by the differential display result.

 The reamplified fragment corresponding to band L36 was subcloned into a cloning vector and sequenced, as described, above, in Section 6.1. Plasmid DNA from
20 four independent transformants was sequenced. All four plasmids were shown to contain the same insert and a consensus sequence of the four sequences was compiled and is shown in FIG. 3A.

 A database search with this consensus sequence
25 resulted in an alignment with greater than 99% identity to a mouse stearoyl-Co-A desaturase gene, SCD1 (Ntambi, J.M. et al., 1988, J. Biol. Chem. 263:17291-17300; Kaestner, K.H. et al., 1989, J. Biol. Chem. 264:14755-14761), which encodes an enzyme that
30 converts ~~saturated~~ fats to mono-unsaturated fats in the liver (FIG. 3B).

 Mouse stearoyl-Co-A-desaturase mRNA is induced in liver upon feeding of fasted animals (Ntambi, J.M. et al., supra; Ntambi, J.M., 1992, J. Biol. Chem. 267:10925-10930). Further, in studies of lean versus
35

obese mice, rats and chickens, stearyl-Co-A enzymatic activity has consistently been reported to be higher in fat than lean animals (Esner, M., 1979, Biochem. J. 180:551-558; Wahle, K.W.J. and Radcliffe, J.D., 1977, Lipids 12:135-139; Legrand, P. et al., 1987, Comp. Biochem. Physiol. 87B:789-792). Additionally, it has been shown that stearyl-Co-A activity is higher in chickens than turkeys (Kouba, M. et al., 1993, Comp. Biochem. Physiol. 105A:359-362). It is considered that turkeys are a low fat animal as compared to chickens. Thus, it is likely that the stearyl-Co-A enzyme is involved in such body weight regulating processes as control, metabolism and storage of dietary components.

Therefore, by utilizing the genetic obesity paradigms described in this Section and in Section 5.1.1.1, above, a differentially regulated gene, the mouse stearyl-Co-A gene, involved in body weight regulation has been identified, thereby corroborating the usefulness of such paradigms in identifying genes important to body weight disorders, and/or body weight or appetite regulation.

7. EXAMPLE: IDENTIFICATION OF GENES
DIFFERENTIALLY EXPRESSED IN
RESPONSE TO SHORT TERM APPETITE
CONTROL PARADIGMS

In the Example presented in this Section, the short term appetite control paradigm, as described, above, in Section 5.1.1.1, is utilized to identify gene sequences which are differentially expressed and which may contribute to body weight disorders and/or may be involved in such processes as body weight regulation or appetite modulation.

7.1. MATERIALS AND METHODS

Short term appetite control paradigm: 45 male C57B1/6J mice 8 weeks of age were received from Jackson labs. The animals were randomized into three groups of 15 mice each, and housed individually on normal mouse chow (West, D.B. et al., 1992, Am. J. Physiol. 262:R1025-R1032)) for 1 week prior to the initiation of the study. Group 1 mice (Control) were maintained on ad lib mouse chow up until the time of sacrifice. Group 2 mice (Fasted) were fasted for 24 hours prior to sacrifice (water continuously available). Group 3 mice (Fasted-Refed) were fasted for 24 hours and then offered a highly palatable meal (mouse chow mixed with peanut butter) for 1 hour prior to sacrifice. All mice were weighed immediately before the initiation of the experiment and immediately afterward. All mice were sacrificed by CO₂ asphyxiation.

RT-PCR analysis: Quantitative RT-PCR was performed as follows. 1-2µg of total RNA, prepared as described, above, in Section 6.1, was reverse transcribed with oligo dT₍₁₂₋₁₈₎ primers and Superscript™ RNAase H-reverse transcriptase (Gibco-BRL, Gaithersburg, MD). Briefly, RNA was combined with 1µl oligo dT (500µg/ml) in a total volume of 11µl. The mixture was heated to 70° C. for 10 minutes and chilled on ice. After a brief centrifugation RNA was reverse transcribed for 1 hour. Aliquots of the first strand cDNA were stored at -20° C. until just prior to use.

Expression levels were determined by PCR amplification of serial dilutions of first strand cDNA. In this procedure, cDNA is serially diluted in water. The dilutions are then batch amplified by PCR using sequence-specific primers. All PCR reactions

are amplified under identical conditions. Therefore, the amount of product generated should reflect the amount of sequence template which was initially present. 5-10 fold dilutions of cDNA were used and enough dilutions were used such that the amount of product subsequently produced ranged from clearly visible, by UV illumination of ethidium bromide-stained gels, to below detection levels. The method described herein can distinguish 10-fold differences in expression levels.

Primers were designed for the amplification of the sequenced amplified bands, which were chosen using the program OLIGO (National Biosciences, Plymouth, MN). All quantitative PCR reactions were carried out in a 9600 Perkin-Elmer PCR machine (Perkin-Elmer). Generally, amplification conditions were as follows: 30-40 cycles consisting of a 95° C. denaturation for 30 seconds, 72° C. extension for 1 minute, 50-60° C. annealing for 30 seconds. Following cycling, reactions were extended for 10 minutes at 72° C.

Other procedures: All other tissue collection, RNA isolation, differential display, sequence analysis, and Northern procedures performed in the experiments described in this Section were as described, above, in Section 6.1.

7.2. RESULTS

Mice, as described, above, in Section 7.1, were utilized as part of short term appetite control paradigms. Briefly, C57B1/6J mice were divided into Control, Fasted, and Fasted-Refed groups, in order to identify genes which are differentially expressed in response to hunger and satiety.

The mice were weighed immediately before the initiation of the study and immediately prior to their sacrifice at the end of the study. Body weights (in grams) were as in Table 3, below:

5

TABLE 3

| | | | |
|------------------------|----------------|---------------|---------------------|
| | <u>Control</u> | <u>Fasted</u> | <u>Fasted-Refed</u> |
| 10 <u>Before Study</u> | 23.9+/-1.3 | 23.3+/-1.1 | 23.2+/-1.4 |
| <u>After Study</u> | 24.4+/-1.4 | 19.3+/-1.1 | 21.7+/-1.4 |

Upon sacrifice, Control, Fasted, and Fasted-Refed tissues were collected and immediately frozen. The tissues collected were: hypothalamus, liver, small intestine, pancreas, stomach, and omental adipose tissue. RNA was collected from the tissue samples obtained and was subjected to differential display, as described, above, in Section 7.1.

20

Utilizing such short term appetite control paradigms and differential display techniques, several gene sequences were identified. Data obtained from such sequences is summarized, below, in Table 4. The differential expression data identifying these gene sequences as corresponding to genes which may be involved in body weight disorders and/or body weight or appetite regulation is listed in the columns headed "Fasted" and "Refed", depending on the paradigm in which differential expression of a given gene was analyzed. Further, the tissue in which the differential expression was observed is noted, as is the difference in expression of each gene in the experimental (either fasted or refed animals) versus control tissues. "↑" indicates that gene expression is increased (i.e., there is an increase in the amount

25

30

35

of detectable mRNA produced by a given gene) in experimental versus control tissue, while "↓" indicates that gene expression is decreased (i.e., there is an decrease in the amount of detectable mRNA produced by a given gene) in experimental versus control tissue. Table 4 also notes whether the gene sequence corresponds to a gene which had previously been identified, and additionally notes the figure in which the nucleotide sequence of the given sequence is listed.

TABLE 4

| Gene | Fasted | Refed | Previously Known | Sequence |
|---------------------------|----------------|-------|------------------|----------|
| P3 (SEQ. ID NO:4, 39) | ↑ Pancreas | | Yes | FIG. 4 |
| P13 (SEQ. ID NO:7, 40-42) | ↑ Pancreas | | Yes | FIG. 5 |
| F5 (SEQ. ID NO:12) | ↑ Adipose | | Yes | FIG. 6 |
| L7/L21 (SEQ. ID NO:18) | ↑ Liver | | Yes | FIG. 9 |
| H27 (SEQ. ID NO:49) | ↓ Hypothalamus | | Yes | FIG. 20 |

In addition to the tissues, listed above in Table 4, in which the initial differential expression was observed, further analysis of the tissue distribution of gene expression of the differentially expressed genes has been conducted. Such an analysis consisted of either Northern or RT-PCR studies, or both, as described, above, in Section 7.1. The tissue

distribution information obtained for the above-listed genes is reported, above, in Table 1 of Section 5.2.1.

Database searches revealed that the genes listed in Table 4, above, identified via the short term appetite control paradigms described herein have previously been identified. Specifically, P3 (SEQ ID NO.:4) represents the gene encoding mouse glutamine synthetase (FIG. 4); P13 (SEQ ID NO.:7, 40-42) represents the gene encoding mouse islet regenerating protein (FIG. 5); F5 (SEQ ID NO.:12) represents the gene encoding mouse α -amylase (FIG. 6); L7/L21 (SEQ ID NO.:18) represents the gene encoding mouse cytochrome c oxidase subunit I (FIG. 9); and H27 (SEQ ID NO:49) represents the gene encoding mouse autoantigen La (FIG. 20).

8. EXAMPLE: IDENTIFICATION OF GENES DIFFERENTIALLY EXPRESSED IN RESPONSE TO GENETIC OBESITY PARADIGMS

In the Example presented in this Section, genetic obesity paradigms, as described, above, in Section 5.1.1.1, were utilized to identify gene sequences which are differentially expressed and which may contribute to body weight disorders and/or may be involved in such processes as body weight regulation or appetite modulation.

8.1. MATERIALS AND METHODS

Genetic obesity paradigms: Animals and animal treatments were as described above, in Section 6.1.

Other procedures: All other tissue collection, RNA isolation, differential display and sequence analysis procedures performed in the experiments

described in this Section were as described, above, in Section 6.1. RT-PCR procedures were as described, above, in Section 7.1.

5 8.2. RESULTS

Ob/ob, db/db, and lean littermate control mice, as described, above, in Section 8.1, were utilized as part of genetic obesity paradigms. The mice were weighed at the end of the study, immediately prior to
10 sacrifice.

Upon sacrifice, tissues were collected from the four groups (i.e., ob/ob, db/db and lean littermate controls) and immediately frozen. The tissues collected were: hypothalamus, liver, small intestine,
15 pancreas, stomach, epididymal or uterine fat pads, and skeletal muscle. RNA was collected from the tissue samples obtained and was subjected to differential display, as described, above, in Section 8.1.

Utilizing such genetic obesity paradigms and
20 differential display techniques, several gene sequences, corresponding to both unique (i.e., previously unknown) and known genes were identified. Data obtained from such sequences is summarized, below, in Table 5. The differential expression data
25 identifying these gene sequences as corresponding to genes which may be involved in body weight disorders and/or body weight or appetite regulation is listed in the columns headed "Ob/ob" and "Pb/db", depending on the paradigm in which differential expression of a
30 given gene was analyzed. Further, the tissue in which the differential expression was observed is noted, as is the difference in expression of each gene in the experimental (either ob or db animals) versus control tissues. "↑" indicates that gene expression is
35 increased (i.e., there is an increase in the amount of

detectable mRNA produced by a given gene) in experimental versus control tissue, while "↓" indicates that gene expression is decreased (i.e., there is an decrease in the amount of detectable mRNA produced by a given gene) in experimental versus control tissue. Further, "+" indicates that gene expression was activated in experimental versus control tissue, i.e., mRNA was detectable in experimental tissue whereas none was detectable in control tissue.

Table 5 also notes whether the gene sequence corresponds to a gene which had previously been identified, and additionally notes in which figure the nucleotide sequence of the given sequence is listed.

TABLE 5

| Gene | <u>ob/ob</u> | <u>db/db</u> | Prev. known | Seq. |
|-------------------------------------|--------------|--------------|----------------|---------|
| F49 (SEQ. ID NO:34) | | +Adipose | No | FIG. 14 |
| murine C5* (SEQ. ID NO:36) | ↑Adipose | ↑Adipose | No | FIG. 16 |
| L31/F74 (SEQ. ID NO:16) | ↓Adipose | ↓Adipose | Yes | FIG. 8 |
| L7/L21 (SEQ. ID NO:18) | ↑Liver | ↑Liver | Yes | FIG. 9 |

| | | | | | |
|----|-------------------------------------|--------|--------|-----|---------|
| 5 | L29 (SEQ. ID NO:20) | ↓Liver | | Yes | FIG. 10 |
| | L38 (SEQ. ID NO:22, 43) | ↑Liver | ↑Liver | Yes | FIG. 11 |
| 10 | L37 (SEQ. ID NO:25, 44-45) | ↑Liver | | Yes | FIG. 12 |
| 15 | L34 (SEQ ID NO:57) | ↑Liver | | Yes | FIG. 22 |

20 The mouse C5 sequence was first identified via sequence homology. C5 was then subsequently tested in ob and db mice, at which time it was identified to represent a differentially expressed gene sequence.

25 In addition to the tissues, listed above in Table 5, in which the initial differential expression was observed, further analysis of the tissue distribution of gene expression of the differentially expressed genes has been conducted. Such an analysis consisted of either Northern or RT-PCR studies, or both, as described, above, in Section 8.1. The tissue distribution information obtained for the above-listed genes is reported, above, in Table 1 of Section 5.2.1.

30 As described above, several of the gene sequences identified via the genetic obesity paradigms of the invention represent previously unknown genes. These include F49 (SEQ ID NO.:34) and C5 (SEQ ID NO.:36).

35

A putative full length cDNA clone (FIG. 14; SEQ ID NO.:34) corresponding to the entire coding sequence of the fat-specific F49 gene has been isolated utilizing the techniques described, above, in Section 5.1.1.2. Hybridization of F49 nucleotide sequences to genomic DNA of several divergent organisms reveals that the F49 gene is conserved in most mammals, including monkeys and humans, while the gene appears to be absent from chicken and yeast.

The F49 coding sequence predicts a 96 amino acid protein (SEQ ID NO.:35), shown in FIG 14. The sequence strongly suggests that the F49 gene product is a secreted protein. Take, for example, the F49 gene product hydropathy plot depicted in FIG. 15. The strongly hydrophobic amino-terminal portion of the amino acid sequence is highly suggestive of a signal sequence characteristic of secreted proteins.

A putative full-length cDNA clone (FIG. 16; SEQ ID NO.:36) corresponding to the entire coding sequence of the murine C5 gene has been isolated utilizing the techniques described, above, in Section 5.1.1.2. The C5 coding sequence predicts the protein whose amino acid sequence is shown in FIG. 16 (SEQ ID NO.:37) This amino acid sequence bears a 50% identity to the mouse brown fat uncoupling protein, and thus represents a newly discovered brown fat uncoupling homologue. In order to further characterize the C5 gene product, a C5 cDNA has been cloned into yeast and mammalian (pcDNA3; Invitrogen) expression vectors to test whether the C5 gene product exhibits uncoupling activity.

Additionally, database searches have revealed that several of the genes identified via the genetic obesity paradigms described herein have previously been identified (see Table 5). For example, L31/F74

(SEQ ID NO.:16) represents the gene encoding the mouse major urinary protein II (FIG. 8); L7/L21 (SEQ ID NO.:18) represents the gene encoding mouse cytochrome c oxidase subunit I (FIG. 9); L29 (SEQ ID NO.:20) represents the gene encoding mouse testosterone 15- α hydroxylase (FIG. 10); L38 (SEQ ID NO.:22, 43) represents the gene encoding mouse 24p3, a lipocalin family member of unknown function (FIG. 11); L37 (SEQ ID NO.:25, 44-45) represents the gene encoding mouse p6-5, a gene which is 86% homologous to rat preproelastase I (FIG. 12); L34 (SEQ ID NO:57; FIG. 22) represents the mouse homolog of a human gene, which will be referred to herein as the "human L34 gene". The human L34 gene, whose sequence can be found in Hou et al. (Hou et al., 1994, In Vitro Cell Dev. Biol. Anim. 30A:111-114, which is incorporated herein by reference in its entirety) was cloned from a hepatoblastoma cell line and encodes a leucine-rich transmembrane protein. The mouse L34 gene exhibits an 82% identity with the human L34 gene at the nucleotide level, and the derived amino acid sequence of the mouse L34 gene product exhibits an 86% identity with the human L34 gene product.

Several of the previously identified genes which these studies have demonstrated to be differentially expressed in obese versus lean control subjects have never before been associated with processes involved in body weight regulation, appetite regulation, or body weight disorders, such as obesity. Among these genes are the genes encoding the mouse major urinary protein II (L31/F74), mouse testosterone 15- α hydroxylase (L29), mouse 24p3 (L38), mouse p6-5 (L37), and human and mouse L34 proteins.

35

9. EXAMPLE: IDENTIFICATION OF GENES
DIFFERENTIALLY EXPRESSED IN
RESPONSE TO SET POINT PARADIGMS

In the Example presented in this Section, set point paradigms, as described, above, in Section 5.1.1.1, were utilized to identify gene sequences which are differentially expressed and which may contribute to body weight disorders and/or may be involved in such processes as body weight regulation or appetite modulation.

9.1. MATERIALS AND METHODS

Set point paradigms: 45 male C57B1/6J mice 8 weeks of age were received from Jackson labs. The animals were randomized into 3 groups of 15 mice each, and housed individually on normal mouse chow for 1 week prior to the initiation of the study. Group 1 mice (Control) were maintained on ad lib mouse chow for an additional five days in order to calculate the daily food intake. Group 2 mice (underweight) then received a fraction of normal food intake (60-90%) so as to reduce and maintain their body weight at approximately 80% of control values. Group 3 mice (overweight) were given a cafeteria diet so as to bring their body weights to 125% of control. The three groups of 15 mice each were then sacrificed by CO₂ euthanasia and tissues were immediately collected. Body weights of the three groups of 15 mice were taken at the time of sacrifice.

Other procedures: All other tissue collection, RNA isolation, differential display, sequence analysis, and Northern procedures performed in the experiments described in this Section were as described, above, in Section 6.1. RT-PCR quantitative

analysis was performed as described, above, in Section 7.1.

9.2. RESULTS

5 Mice, as described, above, in Section 9.1, were utilized as part of set point paradigms. The mice were weighed at the end of the study, immediately prior to sacrifice.

10 Upon sacrifice, tissues were collected from the three groups (i.e., Control, Underweight and Overweight) and immediately frozen. The tissues collected were hypothalamus, liver, small intestine, pancreas, stomach, epididymal fat pads, and skeletal muscle. RNA was collected from the tissue samples
15 obtained and was subjected to differential display, as described, above, in Section 9.1.

Utilizing such set point paradigms and differential display techniques, gene sequences L57, F84 and L31/F74, corresponding to known genes, were
20 identified, as summarized in Table 6, below. In addition to differential expression information, Table 6 also notes in which figure nucleotide sequences of the identified genes are listed.

TABLE 6

25

| Gene | Overweight | Underweight | Previously known | Seq. |
|-------------------------------|------------|-------------|------------------|---------|
| 30 L57 (SEQ. ID NO:29, 46-48) | | ↑ Liver | Yes | FIG. 13 |
| F84 (SEQ. ID NO:52, 54) | | ↑ Adipose | Yes | FIG. 21 |

35

| | | | | |
|-------------------------------|--|--------|-----|--------|
| L31/F74 (SEQ. ID NO:16) | | ↓Liver | Yes | FIG. 8 |
|-------------------------------|--|--------|-----|--------|

5 Database searches have revealed that the L57 (SEQ
ID NO: 29, 46-48) gene sequence corresponds to a
previously known gene, the gene encoding the mouse
orphan nuclear hormone receptor (FIG. 13) and the F84
10 (SEQ ID NO: 53, 55) gene sequence encodes mouse
cytochrome p450 IID (FIG. 21). As discussed, above,
in Section 8.2, the L31/F74 gene sequence encodes the
mouse major urinary protein II. Interestingly, these
gene products have never before been associated with
15 processes involving body weight or appetite regulation
or body weight disorders.

10. EXAMPLE: ISOLATION AND CHARACTERIZATION OF
THE HUMAN C5 GENE.

20 Described in the Example presented in this
Section, is the cloning and characterization of a
human homologue of the mouse C5 gene. The mouse C5
gene is a homolog of the mouse brown fat uncoupling
protein, and, further, as demonstrated, above, in
25 Section 8, the gene is differentially expressed in
obese versus lean littermate controls. Thus, C5 can
represent a gene whose product is involved in body
weight disorders, and/or processes involved in body
weight or appetite regulation. Likewise, human C5 can
30 also represent a gene whose product is involved in
such disorders and/or processes in humans.

10.1. MATERIALS AND METHODS

Human C5 was cloned from a human fetal spleen
library (Stratagene). The probe used for the
35 hybridization was a 0.9 kb partial cDNA clone of the

mouse C5 gene. Filters used for hybridization were NitroPlus 2000 (Micron Separations, Inc.).

Hybridization and washing conditions were as per manufacturer's instructions for low stringency
5 hybridizations, except that the hybridization steps were carried out at 42°C.

Northern analysis was performed as described, above, in Section 7.1.

10 10.2. RESULTS

The mouse C5 gene sequence was used, in conjunction with the methods taught, above, in Sections 5.2.1 and 10.1, to isolate a cDNA clone corresponding to human C5. The nucleotide sequence of
15 the human C5 cDNA clone is shown in FIG. 17 (bottom line; SEQ ID NO:38). The FIG. 17 nucleotide sequence (SEQ ID NO:38) encodes all but the first ten amino acids of the human C5 gene product.

The human C5 amino acid sequence derived from the
20 cDNA sequence is also depicted in FIG. 17 (top line; SEQ ID NO: 56). Using standard well known techniques, the human C5 nucleotide sequence can be used to isolate the most 5' end of the C5 coding sequence. Alternatively, in order to obtain a full length C5
25 protein, a nucleotide sequence encoding the first 10 amino acids of the mouse C5 protein may be ligated to the 5' end of the human cDNA sequence disclosed herein, such that an entire linear C5 coding sequence is generated and a full length C5 protein can be
30 expressed using standard techniques. Further, the entire amino acid sequence of such a C5 protein can be synthesized by, again, utilizing standard techniques well known to those of skill in the art.

The amino acid sequence of such a C5 protein is
35 depicted in FIG. 18 (SEQ ID NO: 51). The first ten

amino acid residues represent sequence derived from the mouse C5 protein while the rest of the amino acid residues represent sequence taken from the human C5 protein. Given the high degree of amino acid conservation which exists between the mouse and the human C5 amino acid sequences, it is expected that the protein depicted in FIG. 18 represents a functional C5 protein.

A Northern analysis of the tissue distribution of human C5 mRNA transcripts was performed, as depicted in FIG. 19. RNA from brain, heart, placenta, lung, liver, muscle, kidney, and pancreas tissues were isolated and analyzed for C5 expression. As in the mouse, the approximately 1.8 kb human C5 transcript exhibits a complex pattern of tissue distribution, with mRNA accumulation appearing to be greatest in muscle tissue.

11. DEPOSIT OF MICROORGANISMS

The following microorganisms were deposited with the Agricultural Research Service Culture Collection (NRRL), Peoria, Illinois, on August 23, 1994 and assigned the indicated accession numbers:

| <u>Microorganism</u> | <u>NRRL Accession No.</u> |
|----------------------|---------------------------|
| famf049a | B-21318 |
| fahs005a | B-21320 |

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention, and functionally equivalent methods and components are within the scope of the invention. Indeed, various modifications of the invention, in addition to those shown and

described herein will become apparent to those skilled
in the art from the foregoing description and
accompanying drawings. Such modifications are
intended to fall within the scope of the appended
5 claims.

10

15

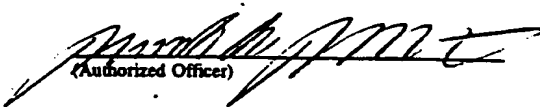
20

25

30

35

International Application No: PCT/ /

| MICROORGANISMS | |
|--|--|
| Optional Sheet in connection with the microorganism referred to on page 127, lines 20-35 of the description * | |
| A. IDENTIFICATION OF DEPOSIT * Further deposits are identified on an additional sheet * | |
| Name of depositary institution * Agricultural Research Culture Collection (NRRL) International Depositary Authority | |
| Address of depositary institution (including postal code and country) * 1815 N. University Street Peoria, IL 61604 US | |
| Date of deposit * <u>August 23, 1994</u> Accession Number * <u>B-21318</u> | |
| B. ADDITIONAL INDICATIONS * (leave blank if not applicable). This information is contained on a separate attached sheet | |
| | |
| C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE * (if the indications are not all designated nations) | |
| | |
| D. SEPARATE FURNISHING OF INDICATIONS * (leave blank if not applicable) | |
| The indications listed below will be submitted to the International Bureau later * (Specify the general nature of the indications e.g., "Accession Number of Deposit") | |
| | |
| E. <input checked="" type="checkbox"/> This sheet was received with the International application when filed (to be checked by the receiving Office) | |
|  (Authorized Officer) | |
| <input type="checkbox"/> The date of receipt (from the applicant) by the International Bureau * | |
| was _____ (Authorized Officer) | |

Form PCT/RO/134 (January 1981)

WO 96/05861

PCT/US95/10918

International Application No: PCT/ /

Form PCT/RO/134 (cont.)

Agricultural Research Culture Collection (NRRL)
International Depositary Authority

1815 N. University Street
Peoria, IL 61604
US

Accession No.

8-21320

Date of Deposit

August 23, 1994

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule which encodes a gene that is differentially expressed in body weight disorders, said gene containing the following nucleotide sequence:
F49 (SEQ ID NO.:34),
C5 (SEQ ID NO.:36), or
human C5 (SEQ ID NO.:38),
or the nucleotide sequence of a gene or gene fragment contained in a cDNA within the following E. coli clone as deposited with the NRRL: famf049a (NRRL Accession No. B-21318), or fahs005a (NRRL Accession No. B-21320).
2. An isolated nucleic acid molecule which encodes a gene that is differentially expressed in body weight disorders, which hybridizes under stringent conditions to the nucleotide sequence of Claim 1 or its complement, or to the gene or gene fragment contained in the cDNA of the E. coli clone of Claim 1 as deposited with the NRRL.
3. An isolated nucleic acid molecule which encodes: an amino acid sequence depicted in FIG. 14 (SEQ ID NO:35), FIG. 16 (SEQ ID NO:37), FIG. 17 (SEQ ID NO:56) or FIG. 18 (SEQ ID NO:51); an amino acid sequence encoded by the nucleotide sequence of Claim 1 or its complement, or an amino acid sequence encoded by the ~~gene or gene fragment contained~~ in the cDNA of the E. coli clone of Claim 1 as deposited with the NRRL.
4. A nucleotide vector containing the nucleotide sequence of Claim 1, 2 or 3.

5. An expression vector containing the nucleotide sequence of Claim 1, 2 or 3 in operative association with a nucleotide regulatory element that controls expression of the nucleotide sequence in a host cell.

6. A genetically engineered host cell containing the nucleotide sequence of Claim 1, 2 or 3.

7. A genetically engineered host cell containing the nucleotide sequence of Claim 1, 2 or 3 in operative association with a nucleotide regulatory element that controls expression of the nucleotide sequence in the host cell.

8. An isolated gene product encoded by the nucleic acid molecule of Claim 1, 2, or 3.

9. The isolated gene product of Claim 8 wherein the gene product is a protein comprising amino acid sequence SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:51 or SEQ ID NO:56.

10. An antibody that immunospecifically binds the gene product of Claim 8.

11. A method for diagnosing body weight disorders, comprising detecting, in a patient sample, a gene transcript or its gene product which is differentially expressed in a body weight disorder relative to a non-body weight disorder.

12. The method of Claim 11 in which the gene is induced in genetically obese individuals.

13. The method of Claim 12 in which the gene encodes stearoyl-CoA-desaturase, cytochrome c oxidase subunit I, a 24p3 protein, a p6-5 protein, a homolog of mouse brown fat uncoupling protein, an F49 protein
5 a C5 protein, or an L34 protein.

14. The method of Claim 11 in which the gene is repressed in genetically obese individuals.

10 15. The method of Claim 14 in which the gene encodes a homologue of major urinary protein II or a testosterone 15- α hydroxylase protein.

16. The method of Claim 11 in which the gene is
15 induced by fasting.

17. The method of Claim 16 in which the gene encodes cytochrome c oxidase subunit 1, glutamine synthetase, islet regenerating protein, or alpha-
20 amylase.

18. The method of Claim 11 in which the gene is repressed by fasting.

25 19. The method of Claim 18 in which the gene encodes an H27 protein or a homologue of rat cytochrome p450 enzyme.

20. The method of Claim 11 in which the gene is
30 repressed by refeeding of a fasted individual.

21. The method of Claim 11 wherein the gene is induced by refeeding of a fasted individual.

35

22. The method of Claim 11 in which the gene is induced in underweight individuals.

23. The method of Claim 22 in which the gene
5 encodes an orphan nuclear hormone receptor protein, a mouse major urinary protein II or an F84 protein.

24. The method of Claim 11 in which the gene is repressed in underweight individuals.
10

25. The method of Claim 11 in which the gene is induced in overweight individuals.

26. The method of Claim 11 in which the gene is repressed in overweight individuals.
15

27. A method for treating a body weight disorder in a mammal, comprising administering to the mammal a compound that modulates the synthesis, expression or
20 activity of a target gene or target gene product so that symptoms of the body weight disorder are ameliorated.

28. The method of Claim 27 in which the compound
25 inhibits the synthesis, expression or activity of the target gene or the target gene product.

29. The method of Claim 27 in which the compound is an antisense or ribozyme molecule that blocks
30 translation of the target gene.

30. The method of Claim 27 in which the compound is complementary to the 5' region of the target gene and blocks transcription via triple helix formation.
35

31. The method of Claim 27 in which the compound is an antibody that neutralizes the activity of the target gene product.

5

10

15

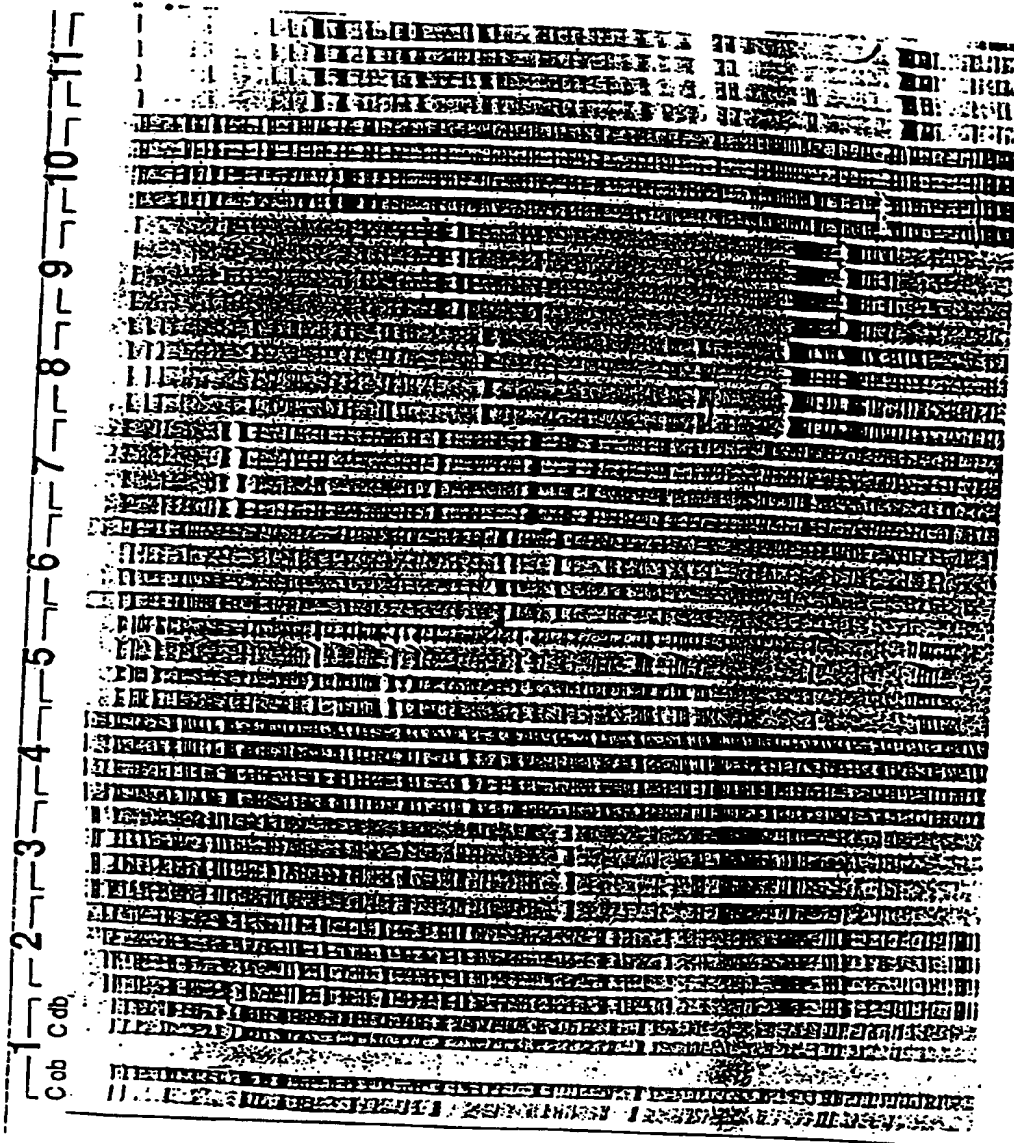
20

25

30

35

1 / 2 2



2 / 2 2

C ob C db



1 2 3 4

FIG. 2

3 / 2 2

10 20 30 40
 GGGCAGTACAACCAGATCCACTTTTATTAGGAACAAATAC 40
 AATCTCAATCAGTACAAGTAGGCTTCAAGAGTTGATATTA 80
 ATGGAAATCATCCAAATTACACTTGGGTCACAAATAATTA 120
 CCCACATAAAAAAGGGAAAAAATACTCATTGAGGGGA 160
 AGGGAAAGGTTTCCTGCAATGGTTTCATGGCAGTGGGTA 200
 210 220 230 240
 GGTAGTCTTGCACTTTGACTGGTCATATCTGTGAGTCTC 240
 TGGGCAGAGCAAA 253 (See ID No. 1)

FIG. 3A

157 CCTGAATGAGATTTTTTTTTTCCCTTTTATGTGGGGTAATTATTTGTGACCCAAGTGTA 98
 |||
 3505 CCTGAATGAGATTTTTTTTTTCCCTTTTATGTGGGGTAATTATTTGTGACCCAAGTGTA 3564
 97 ATTTGGATGATTTCCATTAAATATCAACTCTTGAAGCCTACTTGTACTGATTGAGATTGTA 38
 |||
 3565 ATTTGGATGATTTCCATTAAATATCAACTCTTGAAGCCTACTTGTACTGATTGAGATTGTA 3624
 37 TTGTTCCTAATAAAAGTGGATCTGGTTGTACTGCC 2
 |||
 3625 TTGTTCCTAATAAAAGTGGATCTGGTTGTACTGTC 3660 (See ID No. 2)

217 CAAAGTGCAAGACTACCTACCCACTGCCATGAAAACCATTCAGGAAACCTTTCCCTTCC 158
 |||
 3447 CAAAGTGCAAGACTACCTACCCACTGCCATGAAAACCATTCAGGAAACCTTTCCCTTCC 3506
 157 CCTGAATGAGATTTTTTTTTTCCCTTTTATGTGG 123
 || |||
 3507 TGAATGAGATTTTTTTTTTCCCTTTTATGTGGGG 3541 (See ID No. 3)

FIG. 3B

4 / 2 2

8 CCCTTCCAATACAAGAAGTAAGTGGACTAGACTTCCAGTGATCCCTCTCCCAGCTCTTCC 67
|||||
2347 CCCTTCCAATACAAGAAGTAAGCAGACTAGACTTCCAGTGATCCCTCTCCCAGCTCTTCC 2406

68 CTTTCCCAGTTGTCCCCACTGTAAGTCAAAAG 99 (SEQ ID NO.: 4)
||
2407 CTCTCCCAGTTGTCCCCACTGTAAGTCAAAAG 2438 (SEQ ID NO.: 5)

96 AAAGGATGGAATACCAAGGTCTTTTATTCTCTGTGAAAAAAAAAAAA 143 (SEQ ID NO.: 39)
|||||
2434 AAAGGATGGAATACCAAGGTCTTTTATTCTCTGTGCCAAAAAAAAAGA 2481

(SEQ ID NO.: 6)

FIG. 4

5 / 22

```

318 AGGATACAAAAAATGGGAGGACGAAAACGTGTGAGGCACAGTACTCCTTGGTCTTGAAGTT 377
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3279 AGCATACAAAAAATGGAAGGACGAAAACGTGTGAGGCACAGTACTCCTTTGTCTGCAAGTT 3338

378 CAGAGGCTAA 387 (SEQ ID NO.: 42)
   ||||| |||||
3339 CAGAGCCTAA 3348 (SEQ ID NO.: 11)

```

FIG. 5

6 / 2 2

2 GGCTTAGTCAGCCACTTTACAGACTGGTCTTCTGCTGGCACATACTGTGATGTCATCTC 61
||||| ||||||||||||||||||||||||||||||||||||||||||||
1355 GGCTTTGTCAGCCACTTTACAGACTGGTCTTCTGCTGGCACATACTGTGATGTCATCTC 1414

62 TGGAGATAAGGTCGATGGCAATTGCACTGGACTTAGAGTGAATGTTGGCAGTGATGGCAA 121
||||| ||||||||||||||||||||||||||||||||||||||||||||
1415 TGGAGATAAGGTCGATGGCAATTGCACTGGACTTAGAGTGAATGTTGGCAGTGATGGCAA 1474

122 AGCTCACTTTTCCATTAGTAAGTCTGCTGAGGACCCATTTATTGCAATCCATGCTGACTC 181
||||| ||||||||||||||||||||||||||||||||||||||||||||
1475 AGCTCACTTTTCCATTAGTAAGTCTGCTGAGGACCCATTTATTGCAATCCATGCTGACTC 1534

182 AAAATTGTAAGAATCTATATTAAAGAGATTTGGATTAGGAA 222 (SEQ ID NO: 12)
||||| ||||||||||||||||||||||||||||
1535 AAAATTGTAAGAATCTATATTAAAGAGATTTGGATTAGGAA 1575 (SEQ ID NO: 13)

FIG. 6

7 / 2 2

24 AGAGTCAAGGGCTAGTGGCGCACCGCAGCAGCGCCAGTACCGTGGCGTTCTGGGTACCA 83
||| |||| | | | | | | | | | | | | | | |
247 AGATCCAAGGGGAGTTCCCGATCACCAGCGGCATCAGGTACAAAGGTGTCCTGGGGACAA 306

84 TCCTAACCATGGTGGCACTGAGGGTCCACGCAGCCTCTACAATGGGCTGGTGGCCGGCC 143
|| ||| ||| | | | | | | | | | | | | | | |
307 TCACCACCTTGGCAAAAACGGAAGGGCCCTGAAACTCTACAGCGGTTGCCCGCGGGCC 366

144 TGCAGCGCCAGATGAGCCTTGCCTCCGTCGTCATTGGCCTCTACGACTCTGTCAAACAGT 203
| ||| | | | | | | | | | | | | | | | | | |
367 TCCAGAGACAAATCAGCTTCGCCTCGCTCAGGATCGGCCTCTACGACACGGTGCAGGAGT 426

204 TCTACACCAAGGGCTCAGAGCATGGAGGCATCGGGAGCCGCTCTGGCAGGTAGCAACA 263
||| |||| ||| ||| | | | | | | | | | | | | |
427 TCTTCACCTCGGGGGAAGAAACACCGAGTTTAGGAAGCAAGATCTCGGCCGGCCTAACA 486

264 CAGGTGCCCTGGCCGTGGTTGTAGCCAGCCTACAGATGTGGTAAAGGTCCGCTTCCAGG 323
| | | | | | | | | | | | | | | | | | | | | |
487 CTGGAGGCGTGGCGGTGTTTCATCGGGCAGCCACAGAGGTCTGAAAGTCAGGCTGCAAG 546

324 CTCCAGGCC -332 (SEQ ID NO. : 14)
| | | |
547 CGCAGAGCC 555 (SEQ ID NO. : 15)

FIG. 7

FIG. 8

9 / 2 2

282 TGACCCGTACTTATTACAGCGTACTGCTCCTATTATCACTACCACTGCTAGCCGCAGGC 223
||| |||
5883 TGATCCGTACTTATTACAGCGTACTGCTCCTATTATCACTACCACTGCTAGCCGCAGGC 5942

222 ATTACTATACTACTAACAGACCGCAACCTAAACACAACCTTTCTTTGATCCCGCTGGAGGA 163
||| |||
5943 ATTACTATACTACTAACAGACCGCAACCTAAACACAACCTTTCTTTGATCCCGCTGGAGGA 6002

162 GGGGACCCAATTCTCTACAGCATCTGTTCTGATTCTTTGGGCACCCAGAAGTTTATATT 103
||| |||
6003 GGGGACCCAATTCTCTACAGCATCTGTTCTGATTCTTTGGGCACCCAGAAGTTTATATT 6062

102 CTTATCCTCCCAGGATTGGAATTATTTACATGTAGTTACTTACTACTCCGGAAAAAAA 43
||| |||
6063 CTTATCCTCCCAGGATTGGAATTATTTACATGTAGTTACTTACTACTCCGGAAAAAAA 6122

42 GAACCTTTTCGGCTATATAGGAATAGTATGAAAAAAA 6 (SEQ ID NO. : 18)
||| |||
6123 GAACCTTTTCGGCTATATAGGAATAGTATGAGCAATAA 6159 (SEQ ID NO. : 19)

FIG. 9

10/22

1 TTGTTGCGATCCCAACCTACACTATGAGTTTCTTGTCCTGGCTGCAT 60
|||||
237 TTGTCACGATCCCAACCTACACTATGAGTTTCTTGTCCTGGCTGCAT 296
61 GAGGTTAAAGGAATGATTGAGACCAGACAAGTCAGGGGTTGAACTTAGAAAAGGTCAA 120
|||||
297 GAGGTTAAAGGAATGATTGAGACCAGACAAGTCAGGGGTTGAACTTAGAAAAGGTCAA 356
121 AGGTACAGAAGAAACAGAGGACACTTCGTAGACTTGCAGAGGATATTTCAAAGGTAGCCA 180
|||||
357 AGGTACAGAAGAAACAGAGGACACTTCGTAGACTTGCAGAGGATATTTCAAAGGTAGCCA 416
181 GAGAAGGGGGAAATTATACTATGTTGTCAATAGGAATAATAAAATAATAAAAGTAGATAT 240
|||||
417 GAGAAGGGGGAAATTATACTATGTTGTCAATAGGAATAATAAAATAATAAAAGTAGATAT 476
241 TATTTATGGAA 251 (SER ID NO. : 20)
|||||
477 TATTTATGGCA 487 (SER ID NO. : 21)

FIG. 10

11/22

237 GGTGTTGAGTGTGGCTGACTGGGATGCGCAGAGACCCAATGGTTCAGGCGCTGCCTGTC 178
|| | ||||||||||||||||||||||||||||||||||||||||||||
628 GGGTGGTGTAGTGTGGCTGACTGGGATGCGCAGAGACCCAATGGTTCAGGCGCTGCCTGTC 687

177 TGTCTGCCACTCCATCTTTCTGTGTGCCAGAGAGCCACCTGGCTGCCCCACCAGCCACCA 118
||||||||||||||||||||||||||||||||||||||||||||||
688 TGTCTGCCACTCCATCTTTCTGTGTGCCAGAGAGCCACCTGGCTGCCCCACCAGCCACCA 747

117 TACCAAGGAGCATCTGGAGCCTCTTCTTATTTGGCCAGCACTCCCCATCCACCTGTCTTA 58
||||||||||||||||||||||||||||||||||||||||||||||
748 TACCAAGGAGCATCTGGAGCCTCTTCTTATTTGGCCAGCACTCCCCATCCACCTGTCTTA 807

57 ACACCACCAATGGCGTCCCCTTTCTGCTGAATAAATACATGCCCC 12 (SEQ ID NO.: 22)
||||||||||||||||||||||||||||||||||||||||||
808 ACACCACCAATGGCGTCCCCTTTCTGCTGAATAAATACATGCCCC 853 (SEQ ID NO.: 23)

263 CCGACCAATGCATTGACAACCTGAATGGGTGKT 232 (SEQ ID NO.: 47)
|||||||||||||||||||||||||||
603 CCGACCAATGCATTGACAACCTGAATGGGTGGT 634 (SEQ ID NO.: 24)

FIG. 11

12 / 22

1 TCCGCTCTGGATGCCAGGGTGATTCTGGGGGACCCCTCCACTGCATGGTGAACGGTCAGT 60
|||||
527 TCCGCTCTGGATGCCAGGGTGATTCTGGGGGACCCCTCCACTGCATGGTGAACGGTCAGT 586

61 ATGCTGTCCACGGAGTGACCAGCTTTGTGTCCAGCATGGGCTGTAATGTGCCAGGAAGC 120
|||||
587 ATGCTGTCCAGCGAGTGACCAGCTTTGTGTCCAGCATGGGCTGTAATGTGCCAGGAAGC 646

121 CCACCGTCTTCACCAGAGTCTCTGCTTACATTTTC 155 (SEQ ID NO.: 25)
|||||
647 CCACCGTCTTCACCAGACTCTCTGCTTACATTTCC 681 (SEQ ID NO.: 26)

315 AGCCTTGGGCTCCATCCCTAATACTGCAACAGGAGCAGGGGAATGCTGCTGGTGTCTTGG 374
|||||
839 AGCCTTGGGCTCCATCCCTAATACTGCAACAGGAGCAGGGGAATGCTGCTGGTGTCTTGG 898

375 TATCTGGGGCAAAGGTGGGGGG 396 (SEQ ID NO.: 44)
|||||
899 TATCTGGGGCAAAGGTGGGGGG 920 (SEQ ID NO.: 27)

391 GGGGGGGGTTCATGAAAAGCAACTCAGACTACTGAATCAGATACAGAAAGGCAAATAAAA 450
| |||||
913 GTGGGGGGTTAATGAAAAGCAACTCAGACTACTGAATCAGATACAGAAAGGCAAATAAAA 972

451 ATCAATGTGTTA 462 (SEQ ID NO.: 45)
|||||
973 ATCAATGTGTTA 984 (SEQ ID NO.: 28)

FIG. 12

13 / 22

39 GGGGCAGGGGATCTGCTCAGCTCTATGTTTGAGTTCAGTGAGAAGCTGAATGCCCTCCAG 98
|| |||||
1321 GGAGCAGGGGATCTGCTCAGCTCTATGTTTGAGTTCAGTGAGAAGCTGAATGCCCTCCAG 1380
99 CTCAGTGATGAGGAAATGAGCTTGTTCACAGCAGTTGTTCTGGTATCTGCAG 150 (SER 20 NO: 29)
|| |||||
1381 CTCAGTGATGAGGAAATGAGCTTGTTCACAGCAGTTGTTCTGGTATCTGCAG 1432
(SER 20 NO: 30)

151 GATCGATCTGGAATTGAAAATGTCAACTCAGTGGAGGCTTTGCAGGAAACACTCATCCGT 210
|| |||||
1432 GATCGATCTGGAATTGAAAATGTCAACTCAGTGGAGGCTTTGCAGGAAACACTCATCCGT 1491
211 GCACTAAGGACCTTAATAATGRAAAAACCATCCAAATGAGGCCTCCATTTTACAAAAT 269 (SER 20 NO: 46)
|| |||||
1492 GCACTAAGGACCTTAATAATGRAAAAACCATCCAAATGAGGCCTCCATTTTACAAAAT 1550
(SER 20 NO: 31)

233 AAAAACCATCCAAATGAGGCCTCCATTTTACAAAATTACTTCTAAAGTTGCCAGRTCTT 292
|| |||||
1513 AAAAACCATCCAAATGAGGCCTCCATTTTACAAAATTACTTCTAAAGTTGCCAGATCTT 1572
293 CGATCTTTAAACAACATGCACTC 315 (SER 20 NO: 47)
|| |||||
1573 CGATCTTTAAACAACATGCACTC 1595 (SER 20 NO: 32)

4 GGTAAAGAAGTACAGTGTGGATGACCTGCACTCAATGGG 41 (SER 20 NO: 48)
|| |||||
1285 GGTAAAGAAGTACAGTGTGGATGACCTGCACTCAATGGG 1322 (SER 20 NO: 33)

FIG. 13

14 / 22

1/1. 31/11
 CGC TGT CTC GGG AGG TCT GAA GAC ATG AAG CUG CTT CAG GTT CTC CTT GTT TTC CTC TTT
met lys leu leu gln val leu leu val leu leu phe
 61/21 91/31
 GTG GCA CTT GUA CAT GGT GCA CAG CCC AAA AGA TGT TTT AGC AAC GUA GAA GCC TAC TGT
 val ala leu ala asp gly ala gln pro lys arg cys phe ser asn val glu gly tyr cys
 121/41 151/51
 AAG AAG AAA TGC AGA TTA GTG GAG ATA TCT GAG ATG GGA TGC CTC CAT GGG AAA TAC TGT
 arg lys lys cys arg leu val glu ile ser glu met gly cys leu his gly lys tyr cys
 181/61 211/71
 TGT GGT AAT GAG CTG GAG AAC AAA AAG CAC AAG GAG CAC TCA GTC GTT GAG CAG ACA CTC
 cys val asn glu leu glu asn lys lys his lys glu his ser val val glu glu thr val
 241/81 271/91
 AAA CTC CAA GAC AAG TCA AAA GUA CAA CAC TAT ATG ATC CTG CCC ACG GTC ACA TAC TAC
 lys leu gln asp lys ser lys val glu asp tyr met ile leu pro thr val thr tyr tyr
 301/101 331/111
 ACC ATC ACT ATC TGA ATG AAC CAC TTG TTC ACC AAG CCC GTT GTC CCC TGC AGC CCC ATG
 thr ile ser ile OPN (SEE IO NO. 34)
 361/121 391/131
 GAA TCC AGT CCC CTG CTT CTG TCC TGT CTC TTT CCT TCT GTG AAA CTT CAG TCT GCA CAC
 421/141 451/151
 AAT AAA GTT CGA CCC TTT TGG CTG AAA AAA AAA AAA A (SEE IO NO. 35)

FIG. 14

15 / 2 2

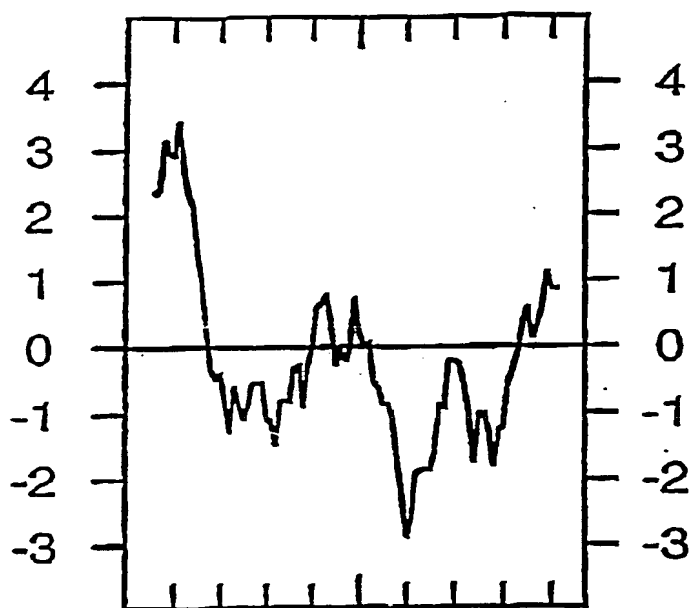


FIG. 15

16/2.2

Start.

| | | |
|---------|---|---|
| 1/1 | atg gct ggt ttc aag gcc aca gat gtg ccc | 31/11 |
| Met | val gly phe lys ala thr asp val pro | cca aca gcc act gtg aag ttc ctg ggg gct |
| 61/21 | | 91/31 |
| | ggg aca gct gcc tgc att gca gat ctc acc | act ttc oct ctg gat acc gcc aag gtc cgg |
| 121/41 | gly thr ala ala cys ile ala asp leu ile | thr phe pro leu asp thr ala lys val arg |
| | | 151/51 |
| | ctg cag atc caa ggg gag agt caa ggg cta | gtg cgc acc gca gcc agc gcc cag tac cgt |
| 181/61 | leu gln ile gln gly glu ser gln gly leu | val arg thr ala ala ser ala gln tyr arg |
| | | 211/71 |
| | ggc gtt ctg ggt acc atc cta acc atg gtg | cgc act gag ggt cca cgc agc ctc tac aat |
| 241/81 | gly val leu gly thr ile leu thr met val | arg thr glu gly pro arg ser leu tyr asn |
| | | 271/91 |
| | ggg ctg gtc gcc ggc ctg cag cgc cag atg | agc ctt gcc toc gtc cgc att ggc ctc tac |
| 301/101 | gly leu val ala gly leu gln arg gln met | ser leu ala ser val arg ile gly leu tyr |
| | | 331/111 |
| | gac tct gtc aaa cag ttc tac acc aag ggc | tca gag cat gga ggc atc ggg agc cgc ctc |
| 361/121 | asp ser val lys gln phe tyr thr lys gly | ser glu his gly gly ile gly ser arg leu |
| | | 391/131 |
| | ctg gca ggt agc acc aca ggt gcc ctg gcc | gtg gtt gta gcc cag oct aca gat gtg gta |
| 421/141 | leu ala gly ser thr thr gly ala leu ala | val val val ala gln pro thr asp val val |
| | | 451/151 |
| | aag gtc cgc ttc cag gct cca gcc cgg gct | ggt ggt ggt cgg aga tac aga gca ctg tog |
| 481/161 | lys val arg phe gln ala pro gly arg ala | gly gly gly arg arg tyr arg ala leu ser |
| | | 511/171 |
| | agc tac aag aac atc acg aga gga ggg atc | cgg ggc ctc tgg aag gga ctc tcc caa tgt |
| 541/181 | ser tyr lys asn ile thr arg gly gly ile | arg gly leu trp lys gly leu ser gln cys |
| | | 571/191 |
| | gcc cgt aat gcc att gtc aac tgt gct gag | ctg gtg acc tat gac ctc atc aaa gat act |
| 601/201 | ala arg asn ala ile val asn cys ala glu | leu val thr tyr asp leu ile lys asp thr |
| | | 631/211 |
| | ctc ctg agc cag ctc atg aca gat gac ctc | oct tgc cag ttc act tct gcc ttc ggg ggc |
| 661/221 | leu leu ser his leu met thr asp-asp leu | pro cys his phe thr ser ala phe gly ala |
| | | 691/231 |
| | ggc ttc tgc acc acc gtc atc gcc toc oct | gtg gat gtg gtc aag acg aga tac atg act |
| 721/241 | gly phe cys thr thr val ile ala ser pro | val asp val val lys thr arg tyr met thr |
| | | 751/251 |
| | ctg ctg ggc cag tac cac agc gca ggt cac | tgt gcc ctt aca tgc tgg gag gag gga ooc |
| 781/261 | leu leu gly gln tyr his ser ala gly his | cys ala leu thr cys ser glu glu gly pro |
| | | 811/271 |
| | ggg ctc ttc aac cag ggg gtt atg oct toc | ttt ctc cgc tgg gga tcc tgg aac gta gtg |
| 841/281 | ala leu phe asn gln gly val met pro ser | phe leu arg leu gly ser trp asn val val |
| | | 871/291 |
| | atg ttt gtc acc tat gag cag ctc caa aga | gcc cta atg gct gcc tac caa tct cgg gag |
| 901/301 | met phe val thr tyr glu gln leu gln arg | ala leu met ala ala tyr gln ser arg glu |
| | | 931/311 |
| | gca cct ttc tga gcc tct cca tgc tga oct | gga ooc tgc ttc cca gcc ctg ccc tgt ctt |
| 961/321 | ala pro phe | OPA (See p. No. 36) |
| | Stop. | 991 |
| | ttt ctc cat oct ctg ooc agt ooc att ctc | ttc oca ttt oct gca ooc cga ttt act toc |
| 1021 | | 1051 |
| | cac ctc acc toc ctg tgc ctc tgt act gat | gac tca cag tga gga ggc ctg aca cca gac |
| 1081 | | 1111 |
| | oct gag ooc tca gcc ctt tct aca gct aag | ooc aca tct tca tct tca toc oca gcc cag |
| 1143 | | 1171/391 |
| | ooc agc cca gct cag cca gcc ttc acc cat | aaa gca agc tca atg tta aaa aaa aaa aaa |
| 1201 | | |
| | aaa aa (See p. No. 37) | |

FIG. 16

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| P | T | A | T | V | K | F | L | G | A | G | T | A | A | C | I | A | 27 |
| CCT | ACT | GCC | ACT | GTG | AAG | TTT | CTT | GGG | GCT | GGC | ACA | GCT | GCC | TGC | ATC | GCA | 60 |
| D | L | I | T | F | P | L | D | T | A | K | V | R | L | Q | I | Q | 47 |
| GAT | CTC | ATC | ACC | TTT | CCT | CTG | GAT | ACT | GCT | AAA | GTC | CGG | TTA | CAG | ATC | CAA | 120 |
| Q | G | P | V | R | A | T | V | S | A | Q | Y | R | G | V | M | G | 67 |
| CAG | GGG | CCA | GTG | CGC | GCT | ACA | GTC | AGC | GCC | CAG | TAC | CGC | GGT | GTG | ATG | GGC | 180 |
| T | M | V | R | T | E | G | P | R | S | L | Y | N | G | L | V | A | 87 |
| ACC | ATG | GTG | CGT | ACT | GAG | GGC | CCC | CGA | AGC | CTC | TAC | AAT | GGG | CTG | GTT | GCC | 240 |
| R | Q | M | S | F | A | S | V | R | I | G | L | Y | D | S | V | K | 107 |
| CGC | CAA | ATG | AGC | TTT | GCC | TCT | GTC | CGC | ATC | GGC | CTG | TAT | GAT | TCT | GTC | AAA | 300 |
| T | K | G | S | E | H | A | S | I | G | S | R | L | L | A | G | S | 127 |
| ACC | AAG | GGC | TCT | GAG | CAT | GCC | AGC | ATT | GGG | AGC | CGC | CTC | CTA | GCA | GGC | AGC | 360 |
| A | L | A | V | A | V | A | Q | P | T | D | V | V | K | V | R | F | 147 |
| GCC | CTG | GCT | GTG | GCT | GTG | GCC | CAG | CCC | ACG | GAT | GTG | GTA | AAG | GTC | CGA | TTC | 420 |
| A | R | A | G | G | G | R | R | Y | Q | S | T | V | N | A | Y | K | 167 |
| CCC | CGG | GCT | GGA | GCT | GGT | CGG | AGA | TAC | CAA | AGC | ACC | GTC | AAT | GCC | TAC | AAG | 480 |
| R | E | E | G | F | R | G | L | W | K | G | T | S | P | N | V | A | 187 |
| CGA | GAG | GAA | GGG | TTC | CGG | GGC | CTC | TGG | AAA | GGG | ACC | TCT | CCC | AAT | GTT | GCT | 540 |
| I | V | N | C | A | E | L | V | T | Y | D | L | I | K | D | A | L | 207 |
| ATT | GTC | AAC | TGT | GCT | GAG | CTG | GTG | ACC | TAT | GAC | CTC | ATC | AAG | GAT | GCC | CTC | 600 |
| N | L | M | T | D | D | L | P | C | H | F | T | S | A | F | G | A | 227 |
| AAC | CTC | ATG | ACA | GAT | GAC | CTC | CCT | TGC | CAC | TTC | ACT | TCT | GCC | TTT | GGG | GCA | 660 |
| T | T | V | I | A | S | P | V | D | V | V | K | T | R | N | M | N | 247 |
| ACC | ACT | GTG | ATC | GCC | TCC | CCT | GTA | GAC | GTG | GTC | AAG | ACG | AGA | AAC | ATG | AAC | 720 |
| G | Q | Y | S | S | A | G | H | C | A | L | T | M | L | Q | K | E | 267 |
| GGC | CAG | TAC | AGT | AGC | GCT | GGC | CAC | TGT | GCC | CTT | ACC | ATG | CTC | CAG | AAG | GAG | 780 |
| A | F | Y | K | G | F | M | P | S | F | L | R | L | G | S | W | D | 287 |
| GCC | TTC | TAC | AAA | GGG | TTC | ATG | CCC | TCC | TTT | CTC | CGC | TTG | GGT | TCC | TGG | GAC | 840 |
| F | G | T | Y | E | Q | L | K | R | A | L | M | A | A | C | T | S | 307 |
| TTC | GGC | ACC | TAT | GAG | CAG | CTG | AAA | CGA | GCC | CTC | ATG | GCT | GCC | TGC | ACT | TCC | 900 |
| P | F | * | (SEQ ID NO: 56) | | | | | | | | | | | | | | 309 |
| CCC | TTC | TGA | GCC | TCT | CCT | GCT | GCT | GAC | CTG | ATC | ACC | TCT | GGC | TTT | GTC | TCT | 960 |
| ATG | CTT | TCC | TTT | TCT | TCC | TTT | CTC | TTT | CCT | CCT | TCC | CTT | CTC | TCC | TTC | CCT | 1020 |
| CCA | CCT | CTT | CCT | TCC | GCT | CCT | TTA | CCT | ACC | ACC | TTC | CCT | CTT | TCT | ACA | TTC | 1080 |
| CAT | TGT | CTC | AGT | GCT | GGT | GGA | GTT | GAC | ATT | TGA | CAG | TGT | GGG | AGG | CCT | CGT | 1140 |
| GAT | CCC | AAG | CGT | CCC | GTC | CCT | TGG | AAA | GTT | CAG | CCA | GAA | TCT | TCG | TCC | TGC | 1200 |
| CCC | AGC | CTA | GCC | CAC | TTG | TCA | TCC | ATA | AAG | CAA | GCT | CAA | CCT | TGA | AAA | AAA | 1260 |

AAAA 1264

(SEQ ID NO: 58)

18/22

FIG. 18

| | | | | | |
|------------|-------------------|-------------|------------|------------|-----|
| MVGFKATDVP | PTATVKFLGA | GTACIADLIT | FPLDTAKVRL | QIQGESQGPV | 50 |
| RATVSAQYRG | VMGTILTMVR | TEGPRSLYNG | LVAGLQRQMS | FASVRIGLYD | 100 |
| SVKQFYTKGS | EHASIGSRLL | AGSTTGALAV | AVAQPTDVVK | VRFQAQARAG | 150 |
| GGRRYQSTVN | AYKTIAREEG | FRGLWKGTSP | NVARNAIVNC | AELVTYDLIK | 200 |
| DALLKANLMT | DDLPCHTSA | FGAGFCTTVI | ASPVDVVKTR | NMNSALGQYS | 250 |
| SAGHCALTML | QKEGPRAFVK | GFMPSEFLRLG | SWDVVMFGTY | EQLKRALMAA | 300 |
| CTSREAPP | (See also No. 51) | | | | 308 |

19 / 22

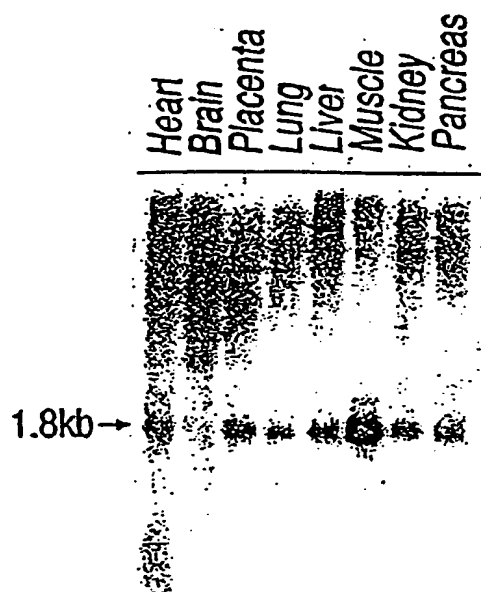


FIG. 19

20/22

FIG. 20

108 GATCCCTGGCCAGAAGTACAAAGACACTAACTGCTAATACTCTTTAAGGAAGATTACTT 167
|||||
647 GATCCCTGGCCAGAAGTACAAAGACACTAACTGCTAATACTCTTTAAGGAAGATTACTT 706
|||||
168 TGCAAAAAAAAAATGAAGAAAGAAAGCAGAGCAAAGTGGAAGCTAAATTAAAAGCTAAACA 227
|||||
707 TGCAAAAAAAAAATGAAGAAAGAAAGCAGAGCAAAGTGGAAGCTAAATTAAAAGCTAAACA 766
|||||
228 AGAGCATGAAGGAAGACACAAGCCAGGAAGTACTGAAACC 267 (SEQ ID NO: 49)
|||||
767 AGAGCATGAAGGAAGACACAAGCCAGGAATACTGAAACC 806 (SEQ ID NO: 50)

21 / 22

FIG. 21

129 CTGTGCTCTGCAAA 117 (SEQ ID NO: 52)
 |||| ||| ||||
 4545 CTGTATCTTCAAA 4557 (SEQ ID NO: 53)

159 TTGGG 155 (SEQ ID NO: 54)
 |||
 4514 TTTGG 4518 (SEQ ID NO: 55)

22/22

1 GGAATTGGGC ACGAGGAAC AGTCTGAGT AGCTCTGIGT
41 GTGTAGACAC TCCCTAAGC ACCCGAGGGC TGCTCTGIGT
61 TGGTGGGGGAG GTAGGAGAGG GGAAGAGGGT TGCTATTCTC
121 ATTCTATACT TTCCCAAGG CTTCOAACCT TATGCAATA
161 ACTCACAATT AAACAGGTCT CAGAACTAAG GGAGCTTTCA
201 ACACTTGGGC TAACICAAGT GCAGTGAGG AAAGGGTTCT
241 CCAAACTGTC TGGGTCTGG CTACCTATGT CACACAGATG
281 AGTGAGTCCA CCGCCCTTC ATTCTCTGAG CTGTGGGTCA
321 CAGCAGGAGT CTTCAGGGC ACTGCTTGG STTGGAAATG
361 TAAGTGGTC CGGAATAGT ATCCCATCA GCTCTCTGGT
401 GHTGTGAGC TCTAGCATGC CATCAGGTT CCAAGGAGGA
441 GAGTACGAA GGTCCAGAG CCGCAAGGTG CACAGAGCAC
481 CCTATAAACC CTGCAGTCT CTTTGAGTCA TGGGGAGGAC
521 TGTGGATCC CTCTTTGAG TCTAAGAGG CTCAGGATC
561 CATATTCTCC ACAAGGTCCA GGTTGGAGT GCATGCATGA
601 AATCAGACTC CAAAAATGA ACCACAAGTC ACAGGAGAGA
641 GAGAGAACTA GTCTGAGCT CGTGGGAAT TGGCAAGAG
681 GTGTAGCTCA GGTAGCCAG GGAGCAATCA AGCTCTGCTT
721 CCTGCTGGC AGCTGTGGT CGCTCTGCC CGAAGTGGT
761 CTGGCCGAGA GGACAGGTT TGCTACAAG ATCTGGGACA
801 AACTTCAGCA GTTAGGTGTC GTATATGATG TCAGTCATTA
841 CAATGCTTCA CTAAAGTAT ATCTTCAAAA TGAATACAAA
881 TTTTCAACCA CTGACTTCT GGCAAGATG GAGGGAGCAA
921 ACATCCAACC AAATGAGTA ACATACCAGA GGCTGATAGC
961 TGCTACTGT AATGTGGGG ACATTGAAGG TGCCAGCAAG
1001 ATCTTGGAT TTATGAAAC GAAAGACCTT CGATCACAG
1041 AGGGGTGTT CAGTCTCTC GTCACAGGC ATGGAGAGC
1081 TGGGGATATG GAAATGCAG AAAATATTCT CACAGTATG
1121 AAACAGGCG GCATGAGCC TGGCCAGAC AGGTATCTGG
1161 CCTGTGTGA TGCATGCT GAGAGGGGTG ACATTGGCCA
1201 GGTAGGCAG ATTCTGGAGA AAGTGGAGAA GTCAGACAT
1241 TACTTCATGG ACGGACTT CTTCAGGTT ATTTTATGCT
1281 TCAGTAAGC TGGCTAACC CACTGAG

FIG. 22.